Establishing a Standard Scientific Guideline for the Evaluation and Adoption of Multi-Tenant Database

Olumuyiwa Oluwafunto Matthew

PhD

August 2016

ESTABLISHING A STANDARD SCIENTIFIC GUIDELINE FOR THE EVALUATION AND ADOPTION OF MULTI-TENANT DATABASE

Olumuyiwa Oluwafunto Matthew

B.Tech, M.Sc

A thesis submitted in partial fulfilment of the
requirements of the University of Wolverhampton for
the degree of Doctor of Philosophy

August 2016

Signature……..
Olumuyiwa O Matthew

Date……..

PUBLICATION LIST

Journal Papers:

Matthew, O., Buckley, K., Garvey, M and Robert Moreton (2016c) MULTI-TENANT DATABASE FRAMEWORK VALIDATION AND IMPLEMENTATION INTO AN EXPERT SYSTEM. *International Journal of Advanced Studies in Computer Science and Engineering* **5**(8).

Matthew, O., Buckley, K and Garvey, M (2016b) STATISTICAL ANALYSIS OF FACTORS THAT INFLUENCES THE EVALUATION AND ADOPTION OF MULTI-TENANT DATABASES International *Journal of Computer Trends and Technology* **37**(2), pp.85-95

Matthew, O., Buckley, K and Garvey, M (2016a) A FRAMEWORK FOR MULTI-TENANT DATABASE ADOPTION BASED ON THE INFLUENCING FACTORS. *International Journal of Information Technology and Computer Science* **8**(3), pp.1-9.

Conference Papers:

Matthew, O., Buckley, K and Garvey, M (2015) *PREDICTING THE IMPACT OF THE FACTORS THAT INFLUENCE THE ADOPTION OF MULTI-TENANT DATABASES*. International Conference on Computer and Information Science and Technology (Cist'15); 05/2015 University of Ottawa, 11-12th May. Canada

Matthew, O., Dudley, C and Moreton, R (2014) A REVIEW OF MULTI-TENANT DATABASE AND FACTORS THAT INFLUENCE ITS ADOPTION: *UKAIS 2014 Conference St Catherine's College,* University of Oxford, 7-9th April. Oxford

ABSTRACT

A Multi-tenant database (MTD) is a way of deploying a Database as a Service (DaaS). A multi-tenant database refers to a principle where a single instance of a Database Management System (DBMS) runs on a server, serving multiple clients organisations (tenants). This technology has helped to discard the large-scale investments in hardware and software resources, in upgrading them regularly and in expensive licences of application software used on in-house hosted database systems. This is gaining momentum with significant increase in the number of organisations ready to take advantage of the technology. The benefits of MTD are potentially enormous but for any organisation to venture into its adoption, there are some salient factors which must be well understood and examined before venturing into the concept. This research examines these factors, different models of MTD, consider the requirements and challenges of implementing MTDs. Investigation of the degree of impact each of these factors has on the adoption of MTD is conducted in this research which focused mainly on public organisations.

The methodology adopted in undertaking this study is a mixed method which involved both qualitative and quantitative research approaches. These strategies are used here to cover statistics (quantifiable data) and experts' knowledge and experiences (abstract data) in order to satisfactorily achieve the aim and objectives and complete the research. Following the involvement of these strategies, a framework was developed and further refined after a second survey was carried out with a quantitative approach. This framework will help prospective tenants to make informed decisions about the adoption of the concept. The research also considers the direction of decisions about MTDs in situations where two or more factors are combined. A new MTD framework is presented that improves the decision making process of MTD adoption.

Also, an Expert System (ES) is developed from the framework which was validated via a survey and analysed with the aid of SPSS software. The findings from the validation indicated that the framework is valuable and suitable for use in practice since majority of respondents accepted the research findings and recommendations for success. Likewise, the ES was validated with majority of participants accepting it and embracing the high level of its friendliness.

The improvement in the acceptance of MTD is also present in this study. This research offers guidelines and recommendations that will assist stakeholders such as Database Service Providers, intending MTD users as well as owners/managers of organisation in resolving the issues surrounding the evaluation and adoption of any MTD model.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AI - Artificial Intelligence

CAPEX - Capital Expenditure

CDB - Container Database

CIST - Conference on Computer and Information Science and Technology

CRM - Customer Relationship Management

CSP - Cloud Service Provider

DaaS - Database as a Service

DBMS - Database Management System

DDL - Data Definition Language

ES - Expert System

GLB - Gramm Leach-Bliley Act

HIPAA - Health Insurance Portability and Accountability Act

HTML - Hyper Text Markup Language

IaaS - Infrastructure as a Service

IDDI - Independent Database Dependent Instances

IDII - Independent Database Independent Instances

ISV - Independent Service Vendor

ITSI - Independent Tables Shared Instances

LAYER - Load As You Query

MMT - Massive Multi-tenant

MTD - Multi -Tenant Database

PaaS - Platform as a Service

PDB - Pluggable Database

RDBMS - Relational Database Management System

RII - Relative Impact Index

ROI - Return on Investment

SaaS - Software as a Service

SLA - Service Level Agreement

SPSS - Statistical Package for the Social Sciences

SQL - Structured Query Language

STSI - Shared Tables Shared Instances

TCO - Total Cost of Ownership

UKAIS - UK Academy for Information Systems

UKOUG - UK Oracle Users Group

UofD - Universe of Discourse

VCE - Virtual Computing Environment

VM - Virtual Machine

XaaS - Everything as a Service

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr. Kevan Buckley (Director of Study - DOS), Dr. Mary Garvey and Prof Robert Moreton for their advice and guidance on my research throughout my studies. They made the process significantly easier than it could have been. I will not forget the effort of my former DOS, Prof Carl Dudley who went on retirement during my first year.

I also would like to thank my family and friends, especially my Father Mr Kolade Matthew for their support and prayers throughout my studies.

My most profound appreciation and love go to my beautiful Wife Mrs Oluwafolakemi Elizabeth Matthew for your love, care, prayer and support.

I dedicate this thesis to my Late Mother Mrs Ifedolapo Matthew.

CHAPTER ONE

1.0 Introduction

Chapter one presents an overview of the research. It describes the research background, scope, motivation, the research aim, research justification, the methodology adopted and significance of the research. The chapter concludes with an outline of the thesis structure. In undertaking any research, it is important to establish the rationale behind such a study as this acts as a reference point against which the research outcomes can be evaluated. Therefore, chapter one serves as an overall introduction to the entire thesis.

1.1 Background of Study.

The use of data is very crucial and important to organisations. Imagine trying to operate a business without knowing who your customers are, what products you are selling, who is working for you, who owes you money, and whom you owe money. All businesses have to keep these types of data and even more. They must have those data available to decision makers when they need them. It can be argued that the ultimate purpose of all business information systems is to help businesses use information as an organisational resource (Coronel and Morris 2016). Therefore, the database system of any organisation plays a very key role in the performance of that organisation. According to Fabling and Sanderson (2016) database is a rich resource for understanding the behaviour and performance of firms. This means that a database system will provide adequate ways of effectively handling the data for any organisation. The efficient data management typically requires the use of a computer database. A database is a shared, integrated computer structure that stores a collection of: end-user data, that is, raw facts of interest to the end user; metadata, or data about data, through which the end-user data are integrated and managed (Coronel and Morris 2016). This is done by a database management system (DBMS). A DBMS is a collection of programs that manages the database structure and controls access to the data stored in the database, in a sense, a database resembles a very well-organised electronic filing cabinet in which powerful software, known as a database management system, helps manage the cabinet's contents (Yaish et al 2013). The collection of computer programs to manage the database is very vital to the optimal performance of that organisation.

The implementation of DBMS for any organisation is based on either the number of users or the database location. The number of users determines whether the database is classified as single-user or multiuser. A single-user database supports only one user at a time while a multiuser database supports multiple users at the same time (Ahn 2014). Location might also be used to classify the database. For example, a database that supports data located at a single site is called a centralized database while a database that supports data distributed across several different sites is called a distributed database (Bell 2014).

The deployment of any of these above mentioned database architecture will require five major components which are data, hardware, software, people and procedures (Tomar and Suruchi 2016). The hardware is the physical devices in the database environment. Operating systems, database management systems and applications make up the software. Examples of people in the database environment are the system administrator, programmers and end users. Procedures are the instructions and rules for the database. There is a huge amount of cost to achieve this for any organisation. This cost is referred to as total cost of ownership (TCO) which is broken down into three major types. These are infrastructural cost, management cost and application development cost (Wang et al 2008 pp94-95). Infrastructural cost includes the cost of hardware, software and utilization costs. Management cost are cost related to the operational activities and processes like lifecycle management, monitoring data backup and restore while application development cost are cost related to meeting each customer additional unique requirement (Wang et al 2008 pp94-95).

However, the advent of database as a service (DaaS) has brought about a huge reduction in cost of deploying DBMS to organisations. Boytsav and Sokolov (2012) argue that DaaS main advantage of its usage is that it takes away the expenditure for deploying infrastructures, eliminate the need for a customer to have his own IT staff and also completely solve the problem of software updating. These costs has been shared among subscribe to the DaaS platform on pay per use basis. DaaS implementation is based on the concept of Multitenant database (MTD), hence the need to study the adoption of the concept multitenant databases.

1.1.1 Concept of Multi-tenancy in Database Systems

Multi-tenancy is a familiar word that comes from the real world of estate building. An estate building will provide multi-tenant housing service to any number of tenants (Banville and Holzel 2012 p2). This could be an individual, couple, family and groups of different sizes. Similarly, a multitenant database is one which provides database supports to a number of

isolated and different groups of users. The concept of multi-tenancy was developed from the service providing technology known as Software as a Service (SaaS). SaaS is a form of cloud computing that involves offering software services in an on-line and on-demand fashion with the internet as the delivery mechanism (Walraven et al 2014). According to Elmore (2011), the concept of a multitenant database has been predominantly used in the context of Software as a Service (SaaS) and Salesforce.com model is often cited as a canonical example of this service paradigm. SaaS is perhaps the most widely recognised of cloud computing service delivery. Software as a Service constitutes a fast-growing business model for the sales of software that bases upon the principle of outsourcing. With SaaS, a service provider hosts an application or software on its infrastructure and delivers it as a service to several organisations. In this model of cloud computing, users share common resources by purchasing access to software and /or data service (Wood and Anderson 2011). The SaaS providers interface with the end users by virtue of provisioning of business application services similar to the ones that have been traditionally self-hosted by the corporate houses (Al-Aqrabi et al 2015 p85). Cloud computing paradigm has emerged to bring large-scale computing, storage resources and data service resources together to build a VCE (virtual computing environment). Cloud computing users can discard the hassles of large-scale investments in hardware and software platforms, in upgrading them regularly and in expensive licenses of application software used to run business processes, related transactions and decision-support systems (Al-Aqrabi et al 2015 p85). An organisation, also referred to as a tenant, subscribes for the service and accesses it across the Internet through standard web technology (Schiller et al 2011 p117). Wood and Anderson (2011) also support this point by saying that tenants or users gain access to SaaS environment which provides access to networks, servers, operating system, storage and application via a cloud service provider (CSP).

A multi-tenant database refers to a principle where single instance of the DBMS runs on a server, serving multiple clients organisation (tenants). Multi-tenant database is one which provides database support to a number of separate and distinct groups of users, also referred to as tenants. A tenant is simply any logically defined group of users that requires access to its own set of data. This definition was substantiated by Bezemer et al (2010 p1) as an architectural pattern in which a single instance of the software is run on the service provider's infrastructure, and multiple tenants access the same instance. This concept provides the ability of a system to provide database management services to different users or customers

without having interference with each other's processes. This reduces effort made in production and the cost incurred in the development.

In a multi-tenant enabled service environment, user requests from different organisations and companies (*tenants*) are served concurrently by one or more hosted application instances and databases based on a scalable, shared hardware and software infrastructure (Gao et al 2011 p324). Such database system must be able to maintain or even increase its performance or efficiency level under larger operational demands.
Multi-tenancy database system is a new technology that can be implemented in both host based and cloud based environment. This research will focus on the trend in this technology, examining the security policies and loopholes in the implementation of the technology and looking at several factors that contribute to the drive towards the multi-tenant database system with a view to producing a standard scientific guideline for the evaluation and adoption of database multi-tenancy.

Multi-tenant data management is a major application of SaaS. Today many companies want to outsource their data to a third party which hosts a multi-tenant database system to provide data management service. Each company is called a tenant. The multi-tenant data management system amortises the cost of hardware, software and professional services to a large number of tenants and thus significantly reduces per- tenant cost by increasing the scale. Thus the multi-tenant database system requires having excellent performance, low-space requirement and good scalability (Ni et al 2012 p2199).

Multi-tenancy is a technique used to consolidate multiple customer applications in a single operational system, is frequently used to obviate the need for separate systems for each tenant (Das et al 2010 p1). The cost and other resources require to have personalised database management system is remove with the adoption of Multi-tenancy. This now means that with multi-tenancy, there is no need for customer having separated on site system for their personal application since such service and resources could be rendered to them through a service provider at a much more reduced cost.

A single instance of the software running on a single server will serve multiple client organisations (tenants), this is a principle also used in software architecture is also referred to as multi-tenancy. Multi tenancy can minimise Hardware/Software costs and human costs per

tenant. Multi-tenant database system has been exploited to store, manage, and retrieve data of tenants. A service provider hosts the multi-tenant database system and each tenant subscribes to the services by doing necessary configuration, loading data to the data center, and then interacts with the services through a standard method, e.g., Web Services. Thus, the cost of ownership of database applications and the maintenance costs are transferred from the individual tenant to the service provider (Ying et al 2011p335). From Pengcheng et al (2015) point of view, clients (tenants) enjoy the desirable features of MTD which include lower upfront investment, pay-as-you-go pricing and reliable performance as specified by the service level agreement (SLA). There is a consolidation at this level which for reduction in operational cost without any revenue loss. This is extremely attractive to database as a service providers. Pengcheng et al (2015) gave examples of MTD as a service providers include salesforce.com's force.com which provides data services in its toolkit for building applications and Amazon's SimpleDB, which provide an API for creating data stores which can in turn, be used for application or pure data storage.

1.2 Research Scope and Motivations

The research scope and motivations for undertaking this study are summarised in the following sections.

1.2.1 Research Scope

The scope of this research focuses on public organisations. The research investigates factors that influence the successful evaluation, adoption and utilization of multi-tenant database concept and further determines the extent to which each of these factors influences the adoption of MTD. This research also help to develop a guidelines and framework that will assist the intending tenants in taking an informed decision about MTD and the level of acceptance of MTD was also looked into in this research.

1.2.2 Research Motivations

This research evolved as a result of the researcher's discontent about the huge amount of investment public sector put into the development of on-premises database system, since many of them still adopt the traditionally on-site system rather than adopting modern technology like MTD. A number of factors which militate against the adoption and use of this new technology were identifies in the research area which include the size of tenant database,

number of tenants, number of users per tenant, growth rate of tenants, growth rate of tenant database, security, cost, time, flexibility, customisation and scalability.

In order to produce such a report, the following objectives were proposed: to review multi-tenant database concept; to investigate existing approaches used in multi-tenant database; to examining the different approaches to determining their advantages and disadvantages; to identify and determining the degree of impact each factor has on decision about MTD adoption; to develop guideline for evaluation and adoption of multi-tenant database; to develop the guideline into a web based expert system; to illustrate practicality of guidelines in a specific public sector organisation and to recommend areas for future research.

1.3 Statement of Research Problem

In recent years, the use of database consolidation is gaining wide acceptance according to Ahmad and Bowman (2011) as a means of reducing the cost and complexity of managing the database system. This shows that a multi-tenant database (MTD) will help reduce the cost and complexity of managing the individual database management system. Gao et al (2011), alludes to the fact that low cost of ownership for companies or tenants and economics of scale is the primary factor in Multi-tenant database adoption. The interest in MTD has been increasing in recent years as illustrated by Hui et al (2009) due to the fact that service providers amortises the cost of hardware, software and professional services to an amount of tenants it serves and therefore significantly reduces per-tenant service subscription fee by use of the economy of scale. This opens avenue for companies or tenants to reduce on the amount spent on the on premise database system.

In spite of the growing number of studies on the concept of MTD and its benefits like in the examining of the requirements for multi-tenant databases and implementation of MDS (Jacobs and Aulbach 2007), schema-mapping techniques (Aulbach et al 2008), a non-intrusive MDS (Gao et al 2011), performance evaluation of the multi-tenant data tier design (Wang et al 2008), Native support of multi-tenancy in RDBMS for SaaS (Schiller et al 2011), Secure and Efficient multitenant database for an Ad hoc cloud (Pippal et al 2011), an elastic multi-tenant database schema for SaaS (Yaish et al 2011) and several other. There is still the need to further this research area into investigating other salient factors that affect the adoption of MTD. Therefore, it is imperative to examine these factors and give recommendations and guidelines to support the adoption process of MTD based on the findings of the research.

1.4 Research Aim

The primary aim of this research is to develop a standard scientific guideline that will support the public sector in their quest to make an informed decision towards the evaluation and adoption multi-tenant database approach putting into consideration several factors that could influence this decision. Likewise, it is also important to determine the degree of impact of these factors on the adoption of MTD.

Keemti (2010); Khan et al (2012) and Yaish et al (2013) has suggested a number of factors which influences MTD adoption but researchers are still not close to arriving at a consensus regarding these factors. Conducting a further research on the adoption of the concept based on these factors would assist in determining steps to follow in the use of the MTD and further help to understand the extent to which these factors impact on the adoption and utilisation of MTD within public sectors. An understanding of these determinants can provide an avenue for policy-makers, stakeholders and practitioners to stimulate the rate of MTD adoption and utilisation within organisations around the world.

1.5 Research Questions.

The development of a research question is a process of looking at an issue that might be a problem and formulating a question about it. Sweet and Grace-Martin (2003) state that research question emphasises a lack or absence of understanding about an issue. It refers to the gap that the researcher intends to address. To achieve the research aim stated in Section 1.4, the following research questions have been formulated which are:

- What are the different methods of implementing a multi-tenant database?
- What level of data privacy and isolation does each of the methods offers?
- What are the factors that influence the decision to adopt a multi-tenant database system?
- What is the degree of influence each of these factors has towards the drive for multi-tenant database with specific reference to the public sector?
- What is the level of acceptance of MTD?

1.6 Research Methodology

The research methodology is the various ways, methods, designs and systems which researchers use in administering and collecting data in a research study (Ezejelue and Ogwu, 1990). It is therefore important to outline the methodology adopted to achieve the above stated characteristics of a good research. This outline is based on the answering the research questions. The research methodology of any study directly impacts the strength and generality of the research (Yang et al 2006).

This section presents an overview of the research methodology applied in this study. The research is exploratory, descriptive and analytical in nature and therefore adopts both quantitative and qualitative research approaches. This involves several issues relating to the concept of Multi-tenant database which are broadly reviewed from several scholar articles. These are not numerically quantifiable and moreover, this process involves experts and their organisations. These are tangible issues such as experience of the experts and operations of the organisations which play a vital role. Therefore, a combination of both strategies are used here to cover statistics (quantifiable data) and experts' knowledge and experiences (abstract data) in order to satisfactorily achieve the aim and objectives and also complete the research.

In this research, the questionnaires were systematically administered in United Kingdom, as a delivery and collection questionnaire under self-administer type to representatives from UK Oracle users group. Also, the questionnaires were administered online as internet and intranet-mediated mode of questionnaire under self-administer type to different set of respondent across all continent of the world which includes: Administrators from different organisations and Users from organisations.

In analysing the first phase of the data collected from the focus group, a quantitative statistical method was adopted known as weighted score method also known as numerical indicator (Abeysekera 2001 p10). This research on MTD is largely quantitative and is concerned with measurement of majorly the nominal and ordinal variables. The data from the survey were coded into SPSS and represented in numerical values. These data were subjected to the following descriptive statistical tools; Percentage Frequency Distribution and Cross Tabulation. A predictive analytical method called Relative Importance Index (RII) was also used on the second phase of data collection to also determine the degree of impact of each factor.

1.7 Significance of study

Multi-tenant Database has increasingly become one of the dominant technologies used by organisations to cut down some capital investment in businesses. In recent times, many organisations are adopting MTD to also increase their return on investment within a short period of time, this is achieved by subscribing to a service provider who will offer to host and manage their expensive data for them. Currently, the area of MTD evaluation, adoption and utilization is still under-researched. Thus, this research is of significant importance, creating an empirically grounded understanding to help intending users, IT managers and service providers in various IT contributions around the world.

The contribution of this research work is three-fold which include: contributions to the general body of knowledge, practical and methodological contributions. In terms of contributions to the general body of knowledge, this research has significant implications for Information Systems (IS) research which seeks to understand and explain issues surrounding MTD models in terms of their adoption and usage. Since this research sets out to investigate the adoption and utilisation of MTD organisational contexts, its findings are aimed at providing a deeper understanding of issues associated with the evaluation, adoption and utilisation of MTD in public sectors. In order words, the research contributes to knowledge by developing an evidence based report that describes the surrounding factors influencing the adoption of MTD and the level of MTD adoption.

On the practical contribution, since this area is still under-researched, results of the study will have significant implications for practitioners, especially in the area of new IT technologies usage in regards to their adoption. It is intended that the quantitative phase of the study will contribute to the statistics of MTD use in organisations around the world whilst the qualitative study will add to the body of literature. Also, the research makes a practical contribution by suggesting ways, through which organisations can successfully evaluate, effectively adopt and utilise MTD in their respective businesses and further contribute towards socio-economic advancement of IT world.

Moreover, the research makes a methodological contribution by employing different data collection techniques (triangulation), in analysing issues and factors relating to MTD adoption, and further examines the extent to which each factor impact on the evaluation and adoption on MTD, especially in conducting their business.

Results from this research will be of great benefit to senior managers, IS executives, business managers, government, amongst others, as the research will help managers to better understand the benefits associated with the adoption and utilisation of MTD by helping to provide a set of guidelines. Findings of the research will also assist to better position stakeholders, researchers and practitioners in their attempts to implement and manage any of the MTD models in their businesses.

Therefore, this research will explore the possibility and viability of using the guideline in a sample public sector in order to bring out the benefits of multi-tenant database and with the intention of changing the perspectives of database and IT users towards the concept.

1.8 Thesis Structure

This thesis is divided into Nine Chapters. Chapter one presents the introduction part to explain the background study, the scope and motivation, the research questions, the aim and objectives of the research.

Chapter Two gives a detailed review of similar technology used before the advent of MTD that is the distributed databases system with its benefits and drawbacks. A detailed review of the importance Multi-tenancy and a description of challenges of MTD implementation, the requirement for MTD. A study of different implementation model and their comparison are reviewed in this Chapter. This Chapter considers previous research in MTD and its uses, and provides background studies about the adoption of MTD and other similar concepts. Part of this Chapter has been presented and published in conference proceeding (Matthew, Dudley, and Moreton, 2014). This Chapter also review the factors that influence the adoption of MTD. These factors are examined and regrouped into four for easy analysis. A section of this Chapter reviews the different regulations in different part of the world and how these could be an influencing factor. Also, the evaluation and adoption of different technology were also reviewed in this Chapter.

Chapter Three discusses the research methodology and evaluates the selection of the research method adopted. The techniques used in this thesis are described. A section in the Chapter describes methods for data collection. The Chapter also outlines the underlying research assumptions that guide computer science related research and also justifies the choice of a quantitative research methodology. The research design, its rationale for the chosen approach and its suitability for this research are all examined.

Chapter Four presents the concept of conceptual framework. An initial framework for MTD evaluation and adoption is developed in this Chapter. The framework was justified based on the hypotheses derived from the literature reviews.

Chapter Five presents the analysis and findings of the first phase of survey which involves the focus group. The second survey analysis is also presented. The Chapter includes background information on the participants involved in both surveys and their level of experience regarding MTD. Part of this Chapter has been presented and published in a conference proceeding (Matthew, Buckley, and Garvey, 2015) and has been published in a journal (Matthew, Buckley, and Garvey, 2016a).

Chapter Six discusses the research findings and relates it to the initial framework. A set of guidelines are developed which leads to the amendments to the initial framework. The comparison of the initial and the new framework is also presented in this Chapter. Part of this Chapter has been published in a journal (Matthew, Buckley and Garvey, 2016b).

Chapter Seven describes the concept expert system and how the framework is developed into an expert system. This Chapter shows the expert system shell adopted and the reasons for its adoption. Part of this Chapter has been published in a journal (Matthew, Buckley, Garvey and Moreton, 2016c).

Chapter Eight describes the validation process and the methodology adopted in the validation procedure, namely external and internal validation. Participants who were involved in the first and second phases of the research were invited to partake in the validation process. They were requested to share their opinions on the research findings and recommendations in a questionnaire survey, in order to validate the proposed framework (Matthew, Buckley, Garvey and Moreton, 2016c).

Chapter Nine summarises the overall findings of the research. The Chapter presents the research outcomes including the achievement of the research questions. Subsequently, the Chapter provides the contributions made by the research, specifically focusing on MTD evaluation and adoption amongst intending users and service providers. The limitations of the research are also presented and finally some areas for further research were identified.

1.9 Summary

Chapter One covers background of the study and has presented the research aim which is aligned with the research topic. The Chapter has also put forward the research questions and has reviewed literature that provides a background to the research. The literature reviewed indicates that there is an increasing demand for research on factors affecting the MTD adoption and to provide a systematic guideline to support intending MTD users in the process of adopting the concept. The background understanding of Multi-tenant database and the problem statement for the research were presented in this Chapter. This Chapter also presented the significance of research by highlighting the proposed contributions of the study. MTD service providers, intending users and the public sectors were highlighted as beneficiaries of the outcome of this study. Finally, the thesis structure was outlined which provided an overview of the entire research and a description of the content contained in the other Chapters. It is intended that the outcome of this research will assist the intending users of MTD to take an informed decision about adopting the concept.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Introduction

This Chapter involves the review of database implementation earlier in place before the advent of MTD, known as distributed database system. Its benefits and drawbacks were also considered in this chapter. Several works that have been carried out by different researchers with regards to multi-tenancy in database were also reviewed. It also identifies different approaches used in the implementation and their requirements. A thorough comparison of these approaches was also carried out here, identifying the pitfalls and benefits of each approach, and the challenges involved in the implementation of multi-tenancy in databases. Some of the things also reviewed are the evaluation and adoption of these different technologies. The concluding part of this Chapter looked at the factors that influence the adoption of the MTD as it is the centre focus of this research.

2.2 Distributed Database Systems

A distributed database (DDB) provides a central database resident on a server that contains database objects. Objects to be replicated are gathered together into distribution packages called “slices,” that are encrypted using a short-lived symmetric key and broken into a succession of short, numbered data packets before being transmitted to client devices (Sutter 2002). Garcia-Molina and Abbott (1987) also said that a distributed database is a collection of named data items, each with an associated value. This means that a distributed database is a database in which portions of the database are stored in multiple physical locations and processing is distributed among multiple database nodes. O'Brien and Marakas (2008) defined distributed database as a database in which storage devices are not all attached to a common processing unit such as the CPU, controlled by a distributed database management system (together sometimes called a distributed database system). It may be stored in multiple computers, located in the same physical location; or may be dispersed over a network of interconnected computers. According to Tomar and Suruchi (2016) a distributed database is a set of several parts that correlate with each other logically over a network of interconnected computers. This means that in a distributed database system, the database is stored and spread physically across computers or sites in different locations that are connected together by some form of data communication network which are spread over

WAN or LAN. The computers may be of different types such as IBM Mainframes, VAXs, SUN work station, PCs etc managed by different operating systems and each fragment of the data base may be managed by a different DBMS such as Oracle, Ingres, and Microsoft SQL server. Unlike parallel systems, in which the processors are tightly coupled and constitute a single database system, a distributed database system consists of loosely-coupled sites that share no physical components. A distributed database (DDB) is a collection of multiple, logically interrelated databases distributed over a computer network. Therefore, a distributed database management system (DDBMS) is the software that manages the DDB and provides an access mechanism that makes this distribution transparent to the users.

According to Rahimi and Haug (2010) in the quest for organisation to manage and access their own data and other workgroup's data, arises the need to share data that are dispersed across an enterprise, a new breed of software to managed dispersed data called distributed database management systems was developed. According to Özsu and Valduriez (2011), there number of things distributed on the platform of DDB which include the processing logic or processing elements, another is the functions which could be delegated to various pieces of hardware or software and also data can be distributed. Data are being used by number of application therefore it might be distributed to a number of processing sites. And finally Özsu and Valduriez (2011) said that the control of execution of various task might also be distributed instead of being performed by one computer system. Abadi (2012) explains the two primary drivers for this trend. First, modern applications require increased data and transactional throughput, which has led to a desire for elastically scalable database systems. Second, the increased globalization and pace of business has led to the requirement to place data near clients who are spread across different locations.

According to Coronel and Morris (2016) there are two processes that ensure the distributed databases remain up-to-date and current: replication and duplication. Replication involves using specialized software that looks for changes in the distributive database. Once the changes have been identified, the replication process makes all the databases look the same. The replication process can be complex and time-consuming depending on the size and number of the distributed databases. This process can also require a lot of time and computer resources. Duplication, on the other hand, has less complexity. It basically identifies one database as a master and then duplicates that database. The duplication process is normally done at a set time after hours. This is to ensure that each distributed location has the same data. In the duplication process, users may change only the master database. This ensures that

local data will not be overwritten. Corbett et al (2013) argued that in an unreplicated distributed database each item has exactly one value. In a fully replicated distributed database an item has n associated values where n is the number of nodes in the system. Each value for a given item is stored at a different node in the system. The values of the item should be the same, but due to updating activity the values may be temporarily different. The values stored at a node constitute the database of that node. A distributed database which is neither unreplicated nor completely replicated is called a partially replicated distributed database. In addition to items and values, a distributed database has a collection of consistency constraints (El Abbadi and Toueg 1986).

Due to demand for system availability and autonomy, and enabled by advances in database and communication technology, distributed database systems are becoming widespread. Many database management systems now support extensions for distribution, and at the same time structurally distributed transaction management systems are available on the market (Ceri and Pelagatti 1984). The designers of distributed database applications are now facing a new and relevant problem: how to distribute the data and programs on different computers to obtain the intended performance, reliability, and availability (Ceri et al 1987). A typical DDBMS architecture is shown below in fig 2.1

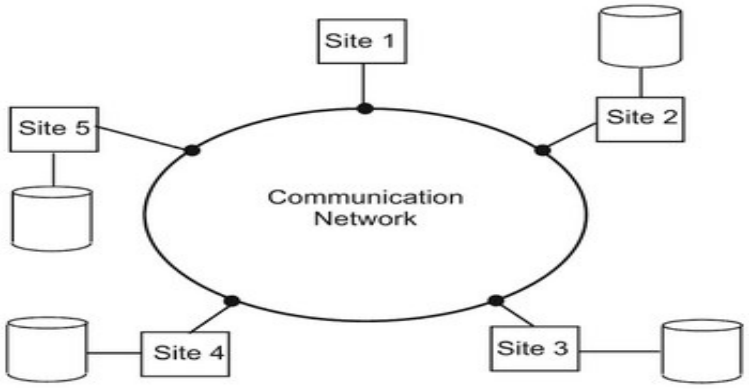


Fig 2.1 – DDBMS Environment (Özsu and Valduriez 2011)

2.2.1 Types of DDBMS

According to Abadi (2012) there are two major types of DDBMS which are classified as:

1. Homogeneous DDBMS.
2. Heterogeneous DDBMS.

In a homogeneous distributed database all sites have identical software and are aware of each other and agree to cooperate in processing user requests. Coronel and Morris (2016) suggested that each location or site must surrenders part of its autonomy in terms of right to change schema or software. In a homogenous distributed database system, all the physical locations have the same underlying hardware and run the same operating systems and database applications. A homogeneous DDBMS appears to the user as a single system. This means that all sites use the same DBMS product. The homogeneous system is much easier to design and manage. Rahimi and Haug (2010) said that some conditions must be satisfied for homogeneous database to occur. These are:

• The operating system used, at each location must be same or compatible.

• The data structures used at each location must be same or compatible.

• The database application (or DBMS) used at each location must be same or compatible.

In a heterogeneous distributed database, different sites may use different schema and software. Difference in schema is a major problem for query processing and transaction processing (Corbett et al 2013). Sites may not be aware of each other and may provide only limited facilities for cooperation in transaction processing. Coronel and Morris (2016) said in heterogeneous systems, different nodes may have different hardware & software and data structures at various nodes or locations which are also incompatible. Different computers and operating systems, database applications or data models may be used at each of the locations. For example, one location may have the latest relational database management technology, while another location may store data using conventional files or old version of database management system. Similarly, one location may have the Windows NT operating system, while another may have UNIX. This now means that since they are not be based on the same underlying data model, it may be that each location is composed of different model like Relational, Network, Hierarchical and Object-oriented DBMSs. Heterogeneous systems are usually used when individual sites use their own hardware and software.

2.2.2 Advantages and Disadvantages of DDBMS

This section looked into the advantages and the disadvantages of the using DDBMS. These are discussed in the sub-section below.

2.2.2.1 Advantages of DDBMS

The advantages of the DDBMS as discussed y different scholar include the following

- **Local autonomy (control)**. According to Tomar and Suruchi (2016) DDMBS has local or site autonomy which make it more flexible and secured. The integrity, storage representation and hardware are controlled locally. At the same time user can access remote data when necessary. This point was also explained by Kaur and Singh (2016) that data accuracy, security and confidentiality are local responsibility while Tomar and Megha (2014) says that local database still works even if the company network is temporarily broken. This shows that the local unit or any of the remote system will still function without reliance on any other system or the central system.
- **No reliance on a central site**. Avoid bottlenecks and system vulnerability. In Tomar and Megha (2014) it was itemized that queries and updates are largely local so that there is no network bottleneck. Since it does not depend on the central, its executions of queries are very fast. Kumar et al (2013) supported this argument by saying that single-site failure does not affect performance of the entire system.
- **Reliability and Availability**. Continue to operate if one or more sites go down or communication links fail. Tomar and Megha (2014) argued that a problem in one part of the organisation will not stop other branches working. Kumar et al (2013) put this advantage as DDBMS as a concept with increased reliability and availability. This is due to the fact that the database is replicated to different location.
- **Speed up of query processing**. Queries about data stored locally are answered faster. Moreover, queries can be split to execute in parallel at different sites or they can be redirected to less busy sites. Kumar et al (2013) argued that the distributed query processing improves the performance of the system. This point was supported by Tomar and Megha (2014) that distributed database is good for the reason that the processing workload is distributed which gives better performance and increased consistency.
- **Modular growth**. It is much easier to add another site than to expand a centralised system. There is the ability of DDBMS to scale up and scale out by adding resources to a single node

and also add more nodes to a system. According to Tomar and Suruchi (2016), scalability is one of the characteristics of DDBMS depending on the application requirements. Distributed database systems scale an increasing data volume by partitioning data across a set of nodes. Resource capacity can be expanded or reduced by horizontal scaling, i.e., adding nodes to the system, or removing nodes from the system respectively (Kuhlenkamp et al 2014).

2.2.2.2 Disadvantages of DDBMS

- **Software complexity and high costs.** A DDBMS that hides the distributed nature from the user and provides an acceptable level of performance, reliability and availability is inherently more complex than a centralized DBMS. According to Tomar and Megha (2014), DDBMS is more complicated to setup and maintain as compared to the central system. Tomar and Suruchi (2016) said even though distributed systems are found in many applications but designing them is a very difficult task, Kumar et al (2013) put this forward by saying that database administrators may have to do extra work to ensure that the distributed nature of the system is transparent and that extra design work must also be done to account for the disconnected nature of the database sometimes. This therefore shows that this extra design effort could be very expensive.
- **Processing overheads:** Increased query processing costs, catalogue management and consistency maintenance. The concurrency control is more difficult in the case of DDBMS because users may access data stored at different replicated copies of original data stored at central site (Tomar and Suruchi 2016). This means that concurrency control mechanism at one site cannot instantaneously know about interactions at other site. Also the overheads could cause a huge processing delay, this is an argument put forward by Tomar and Megha (2014) by saying that distributed databases are not so efficient if there is heavy interaction between the sites. This might cause performance degradation as results of different site requesting to access same set of resources. The allocation of these resources could be a major setback for the whole system. According to Yu and Wang (2011) data or resources allocation is an important issue of DDBMS, because it has significant impact to the efficiency of the whole system.

- **Security and Integrity:** There are so many access points to the system which could hamper the system security and integrity. According to Tomar and Megha (2014), there are many remote entry points to the system compared to a central system which might lead to security

threat. This point was supported by Kumar et al (2013) that remote database fragments must be secured and the infrastructures must also be made secured. This suggests that the multiple accesses to the system can be a source of insecurity to the entire system and subsequently affect the integrity of the system.

2.3 Multi-Tenant Database Technology

With multi-tenancy features, SaaS providers can considerably ease operations and reduce delivery cost with a large number of tenants. As illustrated in Figure 2.2 below, in a multi-tenant enabled service environment, one or more hosted applications instances and database serves the request from different organisations and companies concurrently on a scalable, shared hardware and software infrastructures (Gao et al 2011 p324).

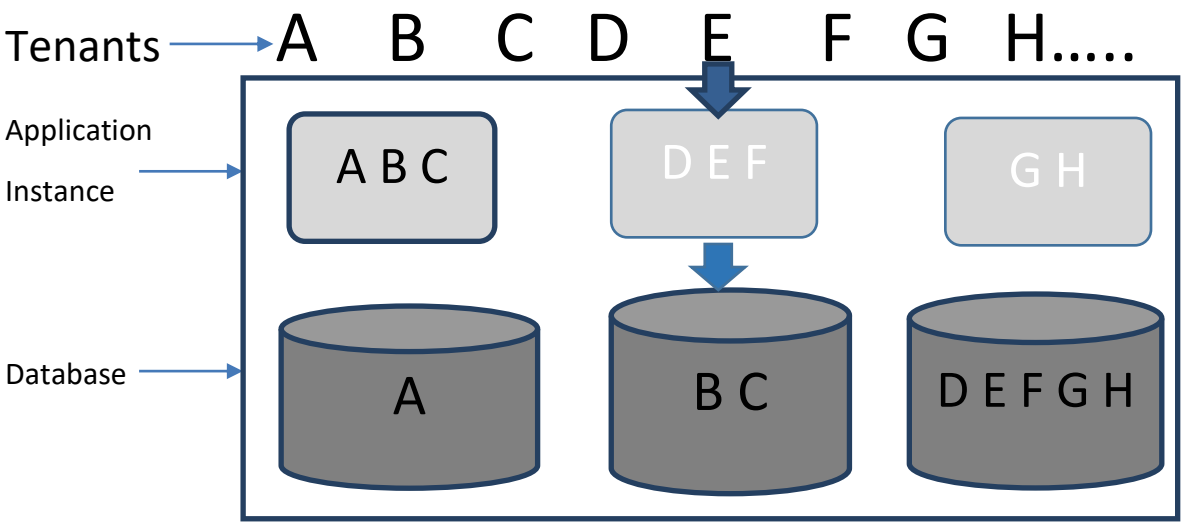


Figure 2.2 - A Multi-Tenant Enabled Service Environment (Gao et al 2011 p324).

Gao et al (2011 p324) focus on the database layer of multi-tenancy as the most important and challenging part of SaaS application looking at cost effective database sharing model and data security isolation among tenants must be guaranteed irrespective of the sharing model. The fact that different tenants with different service level demands and customization needs will further make this multi-tenancy adoption a constraint in practical. Gao et al (2011) came up with a cost-effective, secure, customizable, scalable, and non-intrusive multi-tenant database which will accelerate the migration and development of SaaS.

This Multi-tenancy architecture was also defined by Walraven et al (2014 p2) as an architecture that can be applied at various levels of the software stack: at the infrastructure level (i.e. virtualization), at the OS and middleware level, and even at the application level. This Banville and Holzel (2012 p2) explained multi-tenancy and categorized them based on isolation and sharing of resources. There are four basic classes as shown in figure 2.3: isolated, infrastructure, application and shared tenancy. In isolated tenancy, each tenant has its own instance of the application running with its own instance of database, as well as its own infrastructure to support the deployment (Banville and Holzel 2012 p3). In infrastructure tenancy, each tenant has its own instance of the application running with its own instance of database while they all share the same infrastructure. This kind of multi-tenancy, each tenant has different application and different database but all on the same infrastructure. In application tenancy, each tenant has its own instance of database while at the application and infrastructure level they all run on the same platform.

Finally, shared tenancy has a scenario where all tenants run on the same instance of application, database and infrastructure. This approach is said to be the purest approach of multi-tenancy approaches where everything is been shared (Banville and Holzel 2012 p3).

As we move further to the right side of the figure 2.3, the sharedness of tenancy in each paradigm increases (Banville and Holzel 2012 p3). This means that as you move further to the right side of the diagram, the degree of sharing of the resources increases, taking it to a point where all the required resources are shared.

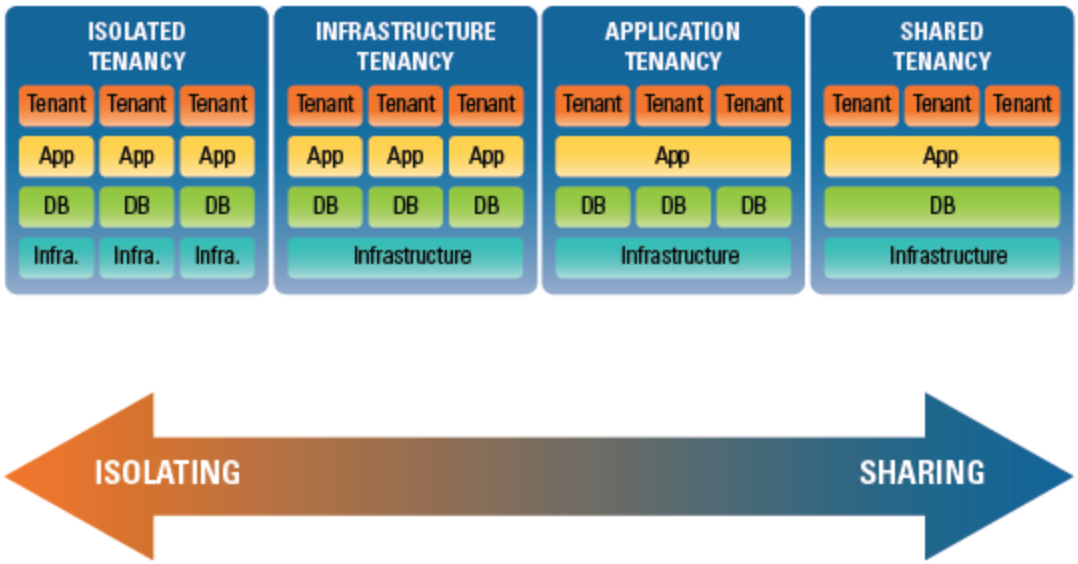


Figure 2.3 - The multi-tenancy continuum (Banville and Holzel 2012 p2)

There are some advantages and disadvantages noted by Banville and Holzel (2012 p3) concerning each of the approaches in figure 2.3. For example, if a tenant requires a customized version of the application and/or the database schema or has special infrastructure requirements, then isolated tenancy is the ideal approach. However, as the number of isolated tenancy deployments increases, the cost of maintaining them also increases. An application provider maintaining an isolated tenancy deployment for an increasing number of tenants will have to invest considerably more time and effort to deploy code changes or carry out database maintenance for all of those tenants.

Also shared tenancy has a comparative reduction in maintenance cost due to one instance of application, database and infrastructure that is shared by all the tenants. It is easy to change code and maintenance is a straightforward and less costly than with isolated tenancy. Isolated tenancy has a high level of scalability due to the ease with which additional tenants can be added to the deployment. While shared tenancy has disadvantages of making it difficult to customize a particular tenant's code, customize schema, or provide any type of specialized service to a particular tenant. Tenants in a shared tenancy configuration may also have increased concerns over data security since their data is stored together on the same database instance (Banville and Holzel 2012 p3).

Walraven et al (2014) focuses on the application-level multi-tenancy which is reported to achieve the highest degree of resource sharing between tenants. It is also Walraven et al that end users from different tenants are simultaneously served by a single application instance on top of the shared infrastructure. There is a crucial disadvantage of inherent limitations in variability when compared with infrastructure-level and middleware –level multi-tenancy. And only requirements that are common to all tenants are satisfied, no support for different and varying requirement of the different tenants. This increase amount of variations and an increasing amount of tenant – specific configuration has led to two essential challenges related to development and customisation of multi-tenant applications identified by Walraven et al (2014 p3). There are;

1. SaaS providers need to be able to manage and reuse the different configurations and software variations in an efficient way, without compromising scalability; e.g. by avoiding additional over-head when provisioning new tenants.

2. Part of realising the scalability benefits of SaaS is achieved by self-service: shifting some of the configuration efforts to the tenant side, e.g. by allowing the tenant to manage his tenant-specific requirements and by automating the run-time configuration process. Therefore, tenants require additional support to manage the configuration in a tenant-driven customisation approach.

2.3.1 Challenges of Implementing Multi-Tenant Database

Ying et al (2011 p335) explained some of the challenges associated with multi-tenancy database development against the traditional database. The first challenge is the data isolation among tenants. Many tenants can share the same database, but the database must ensure the data of these tenants are isolated among each other and no one can get their data other than themselves. The second challenge is to achieve the economics of scales; the database must have the capability of on-demand scale to support large volume of tenants. This means that irrespective of growth in number of tenants and their demand on the database, it must be capable of meeting the demand. Wood and Anderson (2011) argue that complexity through the different and changing demands and requirement of tenant raises further concerns in regards to maintaining and controlling the system. Due to this changing demand over time, the issue relating to scalability and security must be taking into consideration in deploying a multi-tenant environment. The third challenge is to be transparent to current existing application/skill, that is, the cloud developers can easily deploy the existing applications to on multi-tenant database without a large of code change, and the developers can create new multi-tenant application without using new technical knowledge. The fourth challenge is to support different isolations for the same application. This means that the use of different application by different tenant should ensure data isolation to each tenant regardless of the number of tenants involved.

These same constraint was also mentioned by Fang and Tong (2011 p95) that because of the peculiarity need of each tenant, there are problems of: 1) whether the database can afford the increase of both data and request accompanied with the growth of tenants; 2) how the database can meet the specific needs of one tenant efficiently and safely without affecting the others. It seems that the basic challenges associated with this technology since remain the same and different models or approaches were proposed to handle each of these challenges.

2.3.2 Requirement of Multi-Tenant Databases

Multi-tenancy can be applied at the database layer of a hosted service, in order to reduce the high cost of positioning and operating database. Jacob and Aulbach (2007 p2) said that in addition to managing customers' private data, a multi-tenant database should manage customer metadata and shared public data. These metadata should include data like contact information, their location and their features.

The target application for a hosted service will generally have a base schema that specifies all of its standard data. A multi-tenant database should maintain an instance of this base schema for each customer. The unified query language should ensure that DDL statements for modifying the base schema and DML statements for transforming existing data within it are applied to all customers in the farm, within the context of a rolling upgrade. The ability to perform such operations in bulk on the individual databases is essential to minimize downtime during an upgrade (Jacob and Aulbach 2007 p2).

Some of these requirements are also emphasized by Gao et al (2011 pp324-325) which includes:

1. Low Delivery and Operation Cost - Lower the cost of Hardware, software and utility of hosting center (bandwidth, power, space etc.) Lower the cost of human resources to maintain the processes and lifecycle via optimisation and automation.
2. Easy and Low Development Cost - Developers want to focus on business logic development without having to care about multitenant details. Need to use traditional programming model and transform Legacy application.
3. Security isolation - Security Isolation refers to the mechanism where a user is prevented to obtain the priviledge to access resources belonging to other tenants. This is to ensure that each tenant is safeguarded and protected like in the traditional single-tenanted application.
4. Flexibility - The fundamental design point of SaaS is to serve hundreds and thousands of tenants through one or more instance of software application. However, tenant usually has its own requirements such as specific object attribute, business logic etc.

    1. Customisation - Dynamically extend the attributes of existing business objects or create new tenant specific business objects.

2. Diverse SLA- Provide multiple kinds of service options with different SLA (security level, concurrent users, data size, data encryption, backup period etc.).

5. Scalability and availability- Ability to scale very large to support very large customer volume; Incremental scalability, scale-out without impacting the service availability of other tenants.

Nowadays, cloud and SaaS (Software as a Service) have come into our eyeshot. The main idea of SaaS is to use the software as a service and in the using process, several technologies, such as supporting multi-tenant, on-demand use and personalised customisation and so on, are supported. Multi-tenant technology, a technology making integration of multiple applications to one application system, is a good choice for sharing of resources and economies of scale (Sang et al 2012 p179). Cloud in itself is an internet based computing where resources are shared. This service is generally delivered to organisation's computers and / or devices through the internet.

2.3.3 Multi-Tenant Database Implementation

Multi-tenant database architecture is very important for service providers who based on SaaS. This helps in meeting up with the demands of customers or tenants on that platform of the provider. Multi-tenants database architecture is very useful when one instance of database is serving to multiple clients. Only one set of hardware resources is needed to fulfil the requirements of all users. Multi-tenant is based on subscriber model, so user has freedom to avail the facility as per business requirement or can turn off.

Jacobs and Aulbach (2007 p2) identified three different approaches in implementing multi-tenant databases which are: shared machine, shared process, and shared table. These approaches are increasingly better at pooling resources and executing administrative operations in bulk. However, they increasingly break down the isolation between customers, weakening security and increasing contention for resources. Grund et al (2008 p2) made comparisons based on the three approaches of Jacobs and Aulbach (2007) as follows in the shared machine approach each tenants get their own database. The resource sharing is done on machine level. In the shared process approach the tenants share the same physical database process but own different databases. This allows better resource pooling between the tenants but still creates a lot overhead because the schemas need to be maintained separately for all tenants. The last approach is the shared table approach. Using shared tables

the application schema is created once and the different tenants are mapped directly into this schema using different schema mapping techniques.

Schiller et al (2011 p118) further explains these three approaches vary in the degree of consolidation and that the Shared Machine approach allows consolidating only a few tenants onto one machine due to the large main memory footprint of a database instance. The Shared Process approach consumes less main memory per tenant, yet main memory consumption increases quite fast with the number of tenants, as each tenant obtains a dedicated schema instance. In contrast, the main memory consumption of the Shared Table approach remains constant if the number of tenants increases. In Schiller et al (2011 p118) discussion, it was concluded that the shared table approach seems promising for a provider that targets the long tail because it offers the lowest overhead per tenant and, thus, is suitable for a large number of small tenants, e. g. 1,000 tenants each having less than 50 MB of data and at most 5 concurrent users.

Reinwald (2010) in Elmore et al (2011 p2) came out with an improvement on the classification of database multi-tenancy models explored in the past by Jacob and Aulbach (2007), from the three to six classifications. Elmore et al (2011 p2) illustrate that these six classifications were more improved to show different isolation level between tenants possible in all these models. This classification makes selection of a target model interesting and helpful.

Table 2.1: Multitenant database models, how tenants are isolated, and the corresponding cloud computing paradigms. (Elmore et al 2011 p2)

	Sharing Mode	Isolation	IaaS	PaaS	SaaS
1	Shared Hardware	VM	✓	✓	
2	Shared VM	OS User		✓	
3	Shared OS	DB Instance		✓	
4	Shared Instance	Database		✓	
5	Shared database	Scheme		✓	
6	Shared table	Row		✓	✓

Elmore et al (2011) analysed the above table of classification by explaining how the six classifications can be collapse to the more traditional models of multi-tenancy. The models corresponding to rows 1–3 share resources at the level of the same machine with different levels of abstractions, i.e., sharing resources at the machine level using multiple VMs (VM Isolation) or sharing the VM by using different user accounts or different database installations (OS and DB Instance isolation). There is no database resource sharing. Rows 1–3 only share the machine resources and thus correspond to the shared machine model in the traditional classification. Based on Elmore et al (2011) the first three classifications can be collapsed into the traditional shared machine model since all share resources. Elmore et al (2011) also explains that the row 4 and 5 are involve sharing the database process at various isolation levels—from sharing only the installation binary (database isolation), to sharing the database resources such as the logging infrastructure, the buffer pool, etc. (schema isolation), to sharing the same schema and tables (table row level isolation). Rows 4–5 thus span the traditional classes of shared process (for rows 4 and 5). Finally Elmore et al (2011) also explains that row 6 is the shared table model that uses a design which allows for extensible data models to be defined by a tenant with the actual data stored in single shared table. The design often utilizes 'pivot tables' to provide rich database functionality such as indexing and joins (Aulbach et al (2008) in Elmore et al 2011 p2).

Ni et al (2012 p2199) makes a brief comparison of two state-of-the-art approaches which are Independent Tables Shared Instances (ITSI) and Shared Tables Shared Instances (STSI) to design schema. In this comparison ITSI has a poor scalability because there is need for maintenance of very large numbers of tables while STSI achieve very good scalability but with a very poor performance and high space. Ni et al (2014 p2080) further reviewed that there is a third approach called Independent databases Independent instances (IDII). Ni et al (2014) explained that the deployment of IDII is very easy and can be built on top of any current database. It has good data isolation and security but maintenance cost is very expensive. To maintain different database instances, the service provider will have to do much configuration and each instance will allocate a number memory. Also in IDII the number of databases in the server is directly proportional to the number of tenant. This means if there are 100 tenants, then automatically there will be 100 databases instances, thereby making the maintenance cost very high and the scalability is very poor. ITSI has a reduced maintenance cost compared with IDII, thus ITSI has provides a better scalability while STSI

has a better scalability than ITSI because maintenance cost is reduced significantly in STSI than both IDII and ITSI. (Ni et al (2014).
In a multi-tenant setting, the degree of multi-tenancy becomes an additional factor that impacts performance, both for the overall system and the performance that is experienced by each individual tenant. In general, increasing the degree of multi-tenancy decreases per-tenant performance, but reduces the overall operating cost for the DaaS provider. (Lang et al 2012 p702)

Some other approaches for providing multi-tenancy were explained by Pippal et al (2011 p47) which includes; (1) universal table layout, (2) chunk folding, (3) extension tables, (4) pivot tables and (5) multitenant shared table. Each of the approaches has peculiarities which were explained by Pippal et al (2011).

(1) universal table layout

A universal table contains pre specified number of fields [Aulbach et al (2008) in Pippal et al (2011) p47]. It consists of a Tenant_id column, a table column and all the data columns. Tenant_id is used to uniquely identify the data of a tenant whereas the table column refers to the id of the table for that tenant (Pippal et al 2011 p 47). This approach is belief to be easy in implementation and queries are directly applied to the table making queries output faster and in turns improve the system performance. However, this has obvious short comings in that the rows need to be very wide for narrow source tables. Based on this, additional structures must be added to make it feasible (Aulbach et al 2008 p1198).

(2) chunk folding

In chunk folding, it vertically divides the logical tables into chunks and those are folded together into various physical tenants and are joined as needed. One table is used to store the base account information and other table is used to hold the extensions. This approach works by containing the heavily used parts of the schema into base tables and the rest part is mapped into the extensions (Pippal et al 2011 p47). This shows that basic information is on the base table while others are placed in the extension which is link up with the base table.

(3) extension tables

The concept of extension tables came into picture after the development of decomposed storage model described in [Copeland and Khoshafian 2005 in Pippal et al 2011 p47]. It divides a table of n-columns into n 2-column tables that are merged together. One problem with this approach is how to partition the table so that after joining these tables no extra information is generated (Pippal et al 2011 p47). This approach also provides better consolidation but the number of tables increases as the number of tenants increases. As a result of this, there will be wide variety of basic requirements by these different tenants.

(4) pivot tables

In this approach, a pivot table is created for a single column [Grund et al (2008) in Pippal et al (2011) p47]. This table is shared by various tenant's tables. Each pivot table consists of a tenant column, Table column, a "Col" column and a "row" column. Tenant column refers to the particular tenant. Table refers to the particular table for that tenant (Pippal et al 2011 p47). Aulbach et al (2008 p1198) reveals that the Col column specifies the source field a row represent and a single data – bearing column for value of that field.

There are other types mentioned by Aulbach et al (2008 p1197) which were not included in Pippal et al (2011) these include:

1. Basic layout- Aulbach et al (2008 p1197) describes this approach as a very good for its provision of consolidation but there is no extensibility. This involves adding a tenant ID column to each table and share tables among tenants
2. Private table layout - The most basic way to support extensibility is to give each tenant their own private tables. In this approach, the query-transformation layer needs only to rename tables and is very simple. Since the meta-data is entirely managed by the database, there is no overhead for meta-data in the data itself. In this case only moderate consolidation is provided since many tables are required.
   (Aulbach et al 2008 p1197).

2.3.4 Database Migration for Elasticity in Multi-Tenant Database

Multitenant DBMSs often collate multiple tenants' databases on a single server for effective resource sharing. Due to the variability in load, elastic load balancing of tenants' data is critical for performance and cost minimization. On demand migration of tenants' databases to distribute load on an elastic cluster of machines is a critical technology for elastic load balancing. (Das et al 2010 p1). Multi-tenant database (MTD) must be able to handle the varying load requirement of each tenant on the database. This might require adding or sometimes reducing resources depending on whether there is load increase or reduction.

Elasticity is the ability to adapt to varying loads by adding more resources when the load increases, or consolidating the system to lesser number of nodes as the load decreases; all in a live system without disruption in the service (Das et la 2010 p1). Elmore et al (2011 p301) explains that elasticity is the ability to scale up to deal with high load while scaling down in periods of low load. Therefore elastic load balancing is a supreme feature in the design of a modern database management system for a multitenant environment.

Efficient database migration in multitenant databases is an integral component to provide elastic load balancing. Furthermore, considering the scale of the system and the need to minimize the operational cost, the system should be autonomous in dealing with failures and varying load conditions. Migration should therefore be a first class notion in the system having the same stature as scalability, consistency, fault-tolerance, and functionality (Elmore et al 2011 p1). Database migration is a major factor which must be considered in effecting an efficient multitenant database.

However, Luo et al (2015 p87) proposed a novel mechanism called LAYER (Load As You query) to handle the problem associated with excessive workload generated by few most active tenants. This mechanism is met to support multi-tenant applications to server for massive amounts of tenants and at same time be cost efficient. LAYER consolidates into a DBMS a huge number of small tenants with low tenant activeness. LAYER only maintains a moderate number of relevant tables as working tables for answering tenants' queries (Luo et al 2015). When tables involved in a query does not exist in the DBMS and then answers the query based on the restored tables. When an active tenant becomes inactive, the restored tables belonging to the tenant are dropped from DBMS and system resources are reallocated for active tenants.

There is one real-life application scenario of LAYER which is Access 365. And Luo et al presented two implementations of LAYER namely; LAYER – MySQL and LAYER-Volt DB.

2.3.5 Cloud Database Multi-Tenancy

In simple words cloud computing can be defined as getting the work done by sharing and using resources and applications of a network environment without being concerned who is the owner and manager of these resources and applications (Pandithurai et al 2011 p157).
Cloud technology is a means of providing computing facilities without having to spend money to acquire those resources. Cloud gives opportunity for organisation to use expensive resources without having to acquire them on site.

Cloud computing is a new and emerging information technology that changes the way IT architectural solutions are put forward by means of moving towards the theme of virtualization: of data storage, of local networks (infrastructure) as well as software (Bouayad et al 2012 p26). The most widely use definition of cloud is define by Mell and Grance (2010 in Bouayad et al 2012 p26) as a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
Cloud computing is a new technology that provides software and platform as a service to the users on the fly. The users request the resources for some period of time, at certain cost, as decided by the Cloud Service Provider (CSP). Examples of the resources shared include the disk space, processing time, memory utilization, network bandwidth, etc. (Jasti et al 2010 p1).

An increasing number and variety of enterprises are moving workloads to cloud platforms. Whether serving external customers or internal business units, cloud platforms typically allow multiple users, or tenants, to share the same physical server and network infrastructure, as well as use common platform services which include key-value stores, block storage volumes, and SQL databases. These services leverage the expertise of the cloud provider in building, managing, and improving common platforms, and enable the statistical multiplexing of resources between tenants for higher utilisation and cost savings (Shue et al 2012 p16).

Cloud technology makes these expensive resources look as if the resources in on site to the subscriber organisation. Cloud computing is an extremely successful paradigm of service-oriented computing and has revolutionised the way computing infrastructure is abstracted and used. Scalability, elasticity, pay-per-use pricing, and economies of scale are the major reasons for the successful and widespread adoption of cloud infrastructures. Since the majority of cloud applications are data-driven, database management systems (DBMSs) powering these applications are critical components in the cloud software stack (Elmore et al 2011 in Sousa and Machado 2012 p168).
Cloud computing offers many information technology benefits to businesses and organisation and when coupled with cost effectiveness, the complete package is especially attractive to start-up companies or businesses wishing to generate a better cost/benefit approach to use of IT (Flood and Keane 2012 p231).

Cloud computing is an emerging commercial infrastructure model that offers to eliminate the need for companies to maintain in-house high-cost hardware, software, and network infrastructures. It also reduces or even eliminates the high-cost of recruiting technical professionals to support these infrastructures and operate the in-house IT solutions (Hanna et al 2012 p787).

The Cloud Software-as-a-Service (SaaS) where a user can benefit from the Cloud provider's applications that run on the Cloud infrastructure; the Cloud Platform-as-a-Service (PaaS) where a user can deploy applications onto compatible Cloud infrastructures; and the Cloud Infrastructure as a Service (IaaS) where a user can benefit from the computing resources offered and managed by a Cloud service provider (Hanna et al 2012 p787).

There is a more subsuming term called XaaS (Everything as a Service) also used in cloud computing to indicate that any resource could be offered to customers/ tenants by cloud service provider(CSP) either IaaS, or PaaS or SaaS (Banerjee et al 2012 p157).

The number of tenants sharing same SaaS software can scale from dozens to thousands or even more, these tenants may have their particular needs. These scenarios bring two major challenges to traditional relational database: whether the database can afford the increase of both data and request accompanied with the growth of tenants and how can the database meet

the particular needs of one tenant efficiently and safely without affecting the others (Fang and Tong 2011 p95).

2.4 Comparison of Approaches

There are three major approaches identified by Jacobs and Aulbach (2007) which are Shared machine approach, shared process and shared table. Each of these is explained in the subsequent sections.

2.4.1 Shared Machine Approach

This approach offers each tenant a separate database and seems to be the simplest approach towards achieving data isolation. Multiple tenants share the same machine which implies that computing resources and the same data code are also shared among tenants. By giving each tenant a separate database it is quite easy to extend the data model, using standard DDL statements. There is no modification required to the database and no additional isolation considerations that are required under this approach. Data restoration is very easy for all tenants with the use of backups in the case of database system and/or infrastructures failure. This is also probably faster than in the approaches mentioned later. Tenants' migration is quite easy by just moving files from old server to a new server. This is one of the advantages of this approach, while the biggest drawback of this approach is that there is no memory pooling. The first tests done at the Chair for Database Systems at TUM (Jacobs and Aulbach 2007 p3) demonstrate that there is a need of a vast amount of allocated memory for each tested database system to handle just single database instance. This is an indication that this approach cannot scale for more than tens of active customers per server. The costs of hardware are higher with this approach compare to other choice of approaches. The number of tenants that can be on a given database server is limited by the number of databases that the server can simultaneously support. There is a scalability limitation with shared machine approach as connections cannot be shared. There will be hundreds of connections to database system, which subsequently results into a very huge update problem. The update will have to be run for each database which is a great limitation to the hosted services. The CPU is wasted due to hundreds of thousands of queries as results of different update run on different databases and operation speed is slow up. There are replications of multiple data dictionary and the base schema on the same server.

Jacobs and Aulbach (2007 p4) state quite a big issue with one of the most popular commercial database systems. It pre-allocates 200 MB for a database. This could lead to terabytes/megabytes of wasted disk space multiplied across thousands of tenants on the server. This is not the best option for hoisted services but could be considered for businesses with high data isolation requirement. This approach is an option for businesses that are ready to pay extra cost for resources and their maintenance in order to enjoy high level of security and customization.

2.4.2 Shared Process Approach

There is a good balance between scalability, resource pooling and isolation provided by this approach to a great level. Multiple customers share the same database with this approach each customer are offered its own tables. There are two different implementations to this approach. The first type of implementation of this approach is each customer is packed such that the tenant has its own schema and its own table space. This also allows easy migration to new server by simply moving the files to a new server. While the second implementation is to simply attach to the table a unique tenant identifier. Since most of the database systems implement schemas using a lightweight prefixing mechanism this will technically be the same as using an own schema. The sophisticated commercial database systems will then not be able to use the advantage of the separate table space and it will be more difficult to have the good isolation that offers the option with an individual schema. Backup copies can be distributed among the different systems and disks with separate tables' spaces which will results in a better Input and output load balancing. What is more, if this is possible backup copies for an individual tenant can be easily restored. This is a problem with the simple prefixing mechanism since restoring data in the event of failure will not mean simply restoring the most recent backup: this will overwrite the data of every tenant on the same database with backup data, regardless if the tenant has suffered data loss or not. Restoring data to a temporary server will be a solution to this issue, and importing tenant's tables into the production database that are affected. This is time-consuming and can be a major disadvantage.

This approach is considered to be able to scale up to thousands of active customers per server. Jacobs and Aulbach (2007 p4) suggest that this will be a two orders of magnitude improvement over the shared machine approach. Sharing connection pool is possible in this

approach compare to the shared machine approach, and this is because only one database is involved. However, in sharing connection pools, today's database systems have a major disadvantage because all connections will have to be associated with a user having the ability to access everything.

Database vendors' first steps were to allow the user associated with database connection to be selected from the application server since they are aware of this problem. Nevertheless, the current state will mean that security has to be handled in the application layer. This now means that errors in the application code will be accessing private data of other tenants, which is a possible feat for hosted services applications. This might be a major problem for clients with high requirements on security. The issuing of DDL statement to the system can be used to customize tables which will only change the data of the affected tenant. While administrative operations will have the same problems as the shared machine approach.

However, a possible modification of the database system can solve these limitations. The modification of database will give room for keeping a single copy of the data schema and each instance will have to refer to it. There is a great improvement on memory and disk usage The offering of extension tables is another possibility of customizing the data fields. Using this approach would entail joins, which can be quite costly especially when we talk about enormous volume of data. There are propositions and implementations, which offer high performance. Most of the current popular database systems still do not support these features and this can be a serious performance hitch. When it is important that the application be capable of serving an enormous number of tenants with a rather small or few number of servers then the shared process approach is appropriate.

The application has to fulfil also the criteria of having a small number of tables per tenant. Prospective users or tenants should be aware that data isolation will not be as good as it is in the shared machine approach since data will be co-located with that of other tenant, but this result in advantage of lowering the service cost offered to all the tenants. This makes the solution using this implementation particularly beneficial for small or medium-sized businesses.

2.4.3 Shared Table Approach

Shared table approach the data of all tenants is stored in the same tables. The data row has an added column with the tenant identifier. This gives room for actual separation of the data of individual tenants. There must be a specific single value in every query which must specify the column identifying the particular tenant referred to. There is the lowest hardware and backup costs associated with this approach because it allows serving the largest number of tenants per database server compare to other approaches. The only limitation for the number of tenants that can be served using this approach is the number of rows the database can store, which results in significantly better scale up opportunities. There are no challenges sharing the connection to the database. The approach is considerably faster compare to other approaches once the administrative operations and altering of tables can be done in bulk. There are however some bad news when choosing this approach. First of all, additional development effort has to be taken in the area of security, to ensure that tenants can never access other tenants' data, even in the event of unexpected bugs or attacks. Secondly, there is need to execute queries against the database for the purpose of data migration which is a cause for slow system and difficulties of the system. One other major problem is that optimization of queries using indexes is almost not achievable because indexes will have to span the data of all tenants, thus when a customer needs an index on a column, all other customers will have to have the same index.

Integrity constraints and other column-oriented features might also cause problems in this approach. Furthermore, optimization statistics will also have the data from all tenants, making them unrealistic for the single tenant. This problem can be solved only with modification of the database system, so that the tenant identifier is used as a table identifier. Extending a particular table by any tenant can be done in two ways. The first opportunity is by using separate extension tables for each tenant. As already mentioned, this will result in complicated mapping between the logical model the customer sees and the physical model. Extension tables are provided to customers to extend the data model arbitrarily, storing custom data in a separate table and using metadata to define labels and data types for each tenant's custom fields. Therefore an access to a value in the extension table will require joins involving three tables, which is quite a big overhead.

In conclusion, there is the highest scalability and resource pooling in the approach, however, there is the lowest degree of data isolation. This can be used for application which can serve a very large number of tenants with several tables catering for the different tenants on a few numbers of servers.

The summary of all the features identified in the literature review is hereby summarized in Table 2.2 below.

Table 2.2 - Feature of the three models of MTD (Researcher)

Share Machine	**Share Process**	**Share Table**
High Data Isolation	Good data Isolation	Lowest degree of data Isolation
Poor Scalability	Good Scalability	Highest Scalability
High Resource Cost	Good Resource Cost	Very Low Resource Cost
Easy Customization Cost and Effort	Low Customization Cost and Effort	Very High Customization Cost and Effort
No Resource Pooling (Growth)	Low Resource Pooling (Growth)	High Resource Pooling (Growth)

2.5 Evaluation and Adoption of the different Technology

This section is for the literature review on the evaluation and adoption of the different technologies earlier looked into which include the DDBMS, Cloud and MTD. This is explained in the following sub-sections below.

2.5.1 Evaluation and Adoption of DDBMS

The adoption of DDBMS was examined by Gordon and Gordon (1991) by focusing on the issues that influences decision about the adoption of DDBMS. Emphasis was on three major areas which include the organisation structure, organisation culture and top management's attitude towards new technology. This culture was particularly related to its fits with the nature of decision making and with power and control in the organisation. Gordon and Gordon (1991) allude that the selection of DDBMS would mostly be adopted when system is user driven rather than management driven. And that it is also hinged on the fact that its

business needs based rather the management decision driven. Gordon and Gordon (1991) also identified that selection of new information technology was also related to the extent to which decision making was centralised in the organisation. There are two companies investigated in the research which show a strongly centralised decision making organisation, which the owners-founders wanted to retain control of information in the top rather than allowing it to flow down through the organisation. This is an indication that decision towards the adoption of DDBMS can be influence even by the management team of the organisation. This also means that where decision making is decentralised there is likelihood of adoption of a new technology like DDBMS. This proposition was also explained by Lind et al (1989) in Bajaj (2000) that the adoption of a new IT innovation is affected by the organisation size and structure. And Grover and Goslar (1993) in Bajaj (2000) also highlighted that organisational size and centralization affects the adoption new innovations in IT. Gordon and Gordon (1993) conducted another survey to validate the earlier one and it was shown that organisations with decentralized decision making, decentralized IS functions and where top management is favourable to IS were more likely to adopt distributed databases.

In Gordon and Gordon (1993) there were four hypotheses generated, these include; 1. Organisations with decentralized decision making are more likely to adopt DDBMS than those with centralized decision making; 2. Organisations with a decentralized IS structure are more likely to adopt DDBMS than those with a centralized IS structure; 3. Organisations whose top management supports information technology are more likely to adopt DDBMS than those who prefer other solutions; 4. Organisations in which users drive the selection of information technology are more likely to adopt DDBMS than those in which the IS professional staff or top management drives the selection. The result of testing the hypotheses further shows that organisational culture and structure seem to affect the propensity to adopt DDBMS. In particular, this study has suggested that decentralization of decision making, decentralization of the IS function, and management's favourable attitude toward new technology affect the likelihood of DDBMS adoption. IS professionals, IT managers and DDBMS vendors who advocate for the introduction of this new technology must recognize that a set of factors outside their control may affect the likelihood of the adoption which might also complicate the change effort.

Bajaj (2000) conducted a survey using semi-structured interviews with senior IS managers from over 232 firms, and the result shows 30 different identified factors that influences the adoption of a new IS technologies. Bajaj (2000) further mapped these factors into five which

are Software quality, Centralization/distribution, costs, acceptance and backward compatibility.

It is widely accepted in the adoption of innovations literature that complexity and ease of use represent the same concept (Kwon and Zmud 1987; Moore and Benbasat 1991; Tornatzky and Klein 1982). This means that software quality encompasses most of the aspects of complexity or ease of use as identified by other studies.

To centralize IS or not is an issue that has been much debated in the IS literature (e.g., Allen and Boynton 1991; King 1983; Wyner and Malone 1996). The benefits of centralization are widely touted as increased control, uniformity of operations, and economies of scale. The benefits of decentralization include bottom-up productivity improvement, greater autonomy to end users, and the ability to customize IS for frontline organisational functions. This gives more consideration as to consider whether or not centralization would be an important factor in the adoption of future architectures. Any management that seems to want to control the IS facilities and services might discard the adoption of DDBMS. Also in the study of Tomar and Suruchi (2016) it was pointed out that concurrency control and recovery is more difficult issue that is associated with DDBMS because of other many problems that comes with the technique which include dealing with multiple copies of the data items, failure of individual sites, failure of communication links, distributed commit and distributed deadlock. These are some of the vital reasons why DDBMS might not receive a very good attraction.

The costs factor is explicitly mentioned by Tornatzky and Klein (1982) and has been used in several past innovation adoption studies that they list like Gordon and Gordon (1991) using cost as a reason why some management does not want to accept DDBMS. This might be they are trying to avoid spending money on any new technology. This cost is the total cost of ownership as explained earlier in the chapter. This was also buttressed by Kumar et al (2013) that due to increased complexity and a more extensive infrastructure requirement for DDBMS, this will attract more management and labour cost.

The acceptance of the architecture encompasses all aspects of the social approval or image construct (Moore and Benbasat 1991). If a new architecture has greater acceptance with other organisations and the media, then adopting it is likely to lead to greater social approval within and outside the organisation. (Katz and Shapiro 1986; Markus 1987) also allude to the fact that the more accepted the architecture the higher the subsequent adoptions. The popularity

and coupled with the acceptability of a new innovation tends to increase the adoption of that innovation.

Another factor itemized in Bajaj (2000) is the security of the new technology. The new technology must provide minimum acceptable security level for it to be considered by any organisation. Bajaj (2000) said it makes little or no sense to even consider architectures with low security as there is an acceptable level of security required for all architectures. The security level of DDBMS is such an issue identified by other scholars. Tomar and Megha (2014) said that because of the many remote entry points to DDBMS, the security of the system is in great threats. This therefore means that the concept might not have a good adoption rate because of the low level of security it can offer.

2.5.2 Evaluation and Adoption of Cloud

In simple words cloud computing can be defined as getting the work done by sharing and using resources and applications of a network environment without being concerned who is the owner and manager of these resources and applications (Pandithurai et al 2011 p157). Cloud technology is a means of providing computing facilities without having to spend money to acquire those resources. Cloud gives opportunity for organisation to use expensive resources without having to acquire them on site. According to Kim et al (2009) there are three types of potential users of the cloud comping services which include consumers, small organisation medium and large organisations. The requirement for each type of users differs depending on the services needed by these users. Kim et al (2009) said the Consumers and small organisations have relatively simpler requirements for adopting a new technology than medium to large organisations, and have much less to lose if the adoption goes awry. Despite the huge benefits of cloud computing, it also obvious that there are some issues regarding the adoption of cloud computing. Kim (2009) itemizes several concerns that users have regarding the adoption of cloud computing which include availability, security and privacy, support, interoperability, and compliance. Other issues were identified in addition to the above by Kim et al (2009) which are private clouds, integration, cost, and environment. These points were also buttressed by Avram (2014) where some barriers to cloud adoption were listed as follows; security and privacy, connectivity and open access, reliability, interoperability, economic value, change in the IT organisation and political issues due to global boundaries. These are issues one need to examine before venturing into the concept of cloud services

because these concerns are as results of users' data, applications, and computing resources will no longer be under their control. According to Kim et al (2009) among all these, availability, security, and performance are quality of service (QoS) issues. Only some of the adoption issues matter to consumers and small organisations while all of them are of concern to medium to large organisations. This shows that all these issues are very important to the medium and large scale organisations.

There are other studies that recently address the adoption issues that surround the decision of adopting cloud computing in the industrial revolution in terms of implications for new technological innovation. Kshetri (2013) itemises issues such as security, privacy and availability are among the topmost concerns in organisations' cloud adoption decisions rather than the total cost of ownership. According to Kshetri, the primary concerns related to security, privacy and confidentiality might generate huge cost implication that might outweigh the benefits of this concept. Organisations worry about hidden costs associated with security breaches or lawsuits tied to data breach. Businesses and consumers are cautious in using it to store high-value or sensitive data and information (Goodburn & Hill, 2011) in Kshetri (2013).

In Lin and Chen (2012) survey with 19 Taiwanese IT professionals, it was discovered that though the benefits of cloud computing such as its computational power and ability to help companies save cost. It was also agreed that there are primary concerns that IT managers and software engineers have with the compatibility of cloud with companies' policy, IS development environment, and business needs. This will be another form of issues in the adoption of Cloud which according to Lin and Chen drives investor away from the technology. Cloud adoption decisions are challenging because of a range of technical and socio-technical factors which include concern about cost, confidentiality and control (Khajeh-Hosseini 2012). Khajeh-Hosseini said it is unlikely that all organisations will completely outsource their back-end computing requirements to a cloud service provider. Therefore, tenants might not like to adopt cloud based on these factors. Khajeh-Hosseini concluded by listing the challenges that needed to be addressed as follows: (i) to provide accurate information on costs of cloud adoption; (ii) to support risk management; and (iii) to ensure that decision makers can make informed trade-offs between the benefits and risks.

Yeboah-Boateny and Essandoh (2014) conducted a study to look at the influencing factors towards the adoption of cloud. This study looked at this from two different perspectives

which are drivers and barriers of cloud computing. The barriers inhibiting the adoption of cloud computing in Yeboah-Boateny and Essandoh study include the following: lack of internal expertise and knowledge, poor internet access and connectivity, security and data privacy, lack of trust, integration with in-house and existing systems, loss of control, differences in international statutory laws and regulations, delay in the transfer and migration of data, lack of confidence in ability and promise of the cloud and finally the lack of standards. This study used relative importance index to determine the level of importance of each of these factors. The result of this analysis shows that some of the factors have high level of importance which include; lack of internal expertise and knowledge, poor internet access and connectivity, security and data privacy, lack of trust, integration with in-house and existing systems. The rest of the factors have a medium level of importance. Also in Gangwar et al (2015) study some concern and challenges associated with the adoption of cloud computing were identified such as security, service availability, performance, higher costs (when compared to on-premise implementation) associated with the subscription model, lack of interoperability standards, difficult integration with on-premise applications and limited customization facilities. These factors are similar in all the literatures, pointing to the same set of challenges which needed to be considered before adopting cloud technology.

2.5.3 Evaluation and Adoption of MTD

Based on all the identified literature so far, the evaluation and adoption of MTD is an aspect of IT research that seems not to have been tapped or still at elementary stage. Though, the concept is somehow linked to an aspect of the cloud technology. As cloud computing has risen across the enterprise landscape, a new form of information delivery has taken hold. Commonly referred to as Database as a Service, or DaaS, a platform that is based on MTD technology. Gessert et al (2014) supported this argument by saying that there has been a popular shift in application design towards relying on DaaS systems to manage application data where the cloud database service takes the place of a classic application server and allows applications (in particular mobile and web applications) to directly connect to it. This approach promises to finally crack the puzzle that has inhibited enterprise data shops for years i.e the challenge of sharing access to siloed data stores. A survey of more than 300 Database Administrators (DBAs) and IT professionals finds growing interest in DaaS as a viable approach to serving their enterprise's need for greater agility and faster time to market with cloud computing. The survey was conducted by Unisphere Research, a division of

Information Today, Inc in partnership with Oracle among members of the Independent Oracle Users Group (IOUG) (McKendrick 2016). Highlights of the research McKendrick (2016) include the following findings:

1. Database as a Service (DaaS) is taking off, with adoption tripling over the next 24 months. There will be a significant amount of enterprise data shifting to the cloud over the next 24 months as well, as enterprises re-think data management in the cloud. The survey finds an impending increase in organisations running significant portions of the workloads (defined as greater than 25%) in the cloud—from a total of 14% today to 43% within the next 2 years. The gradual shift to database clouds is already underway for a number of organisations with a notable jump from the current 9% level to 25% within the next 2 years.
2. Enterprise customers see a future with hybrid approaches, relying on a combination of private and public cloud resources. As the number of cloud services within enterprises grows, there will be more reliance on both for cost mitigation, as well as backup and continuity.
3. Cloud is increasingly seen as a highly agile and robust platform for enterprise application development. More development work is going to the cloud. Database backup and disaster recovery, along with enterprise applications are the areas seeing the greatest returns from cloud implementations.
4. Significant segments of Oracle Database shops are adopting a range of technologies to move their DaaS efforts forward into their enterprises. Oracle Multitenant, Oracle Enterprise Manager, and Exadata are becoming mainstays for DBAs and professionals seeking to deliver information on-demand to whomever and wherever it is needed.

These above findings about DaaS were also in agreement with the findings from 451 Research (2016) which are as follows:

1. DaaS which is based on MTD technology is in the very early stages of adoption, with most data-related workloads deployed on-premises.
2. The key drivers for cloud databases are cost avoidance, flexibility, IT rejuvenation, transformational change, and data gravity.
3. One company interviewed estimated that it can save 60-70% in terms of licensing costs, and maintenance and upgrades by adopting cloud databases.

4. Interviewed companies also cited operational efficiencies: reduced database administration requirements, lower server configuration and management overheads, and shorter development cycles.
5. The primary barriers to cloud database adoption are security issues (both perceived and real) liability, performance, cost and people and process change.
6. Companies interviewed expect security concerns to abate over time, and already have plans to deploy more mission critical applications involving sensitive data to cloud databases.

Though the findings of these scholars are similar with little or no differences in them, but it is obvious that the adoption of MTD is still in its early stage. The above findings need more verification and validation for it to be generalised. Hence the need for this research which is meant to extend further for contributions and opportunity to create more avenues for further studies.

2.6 FACTORS INFLUENCING THE CHOICE OF DATABASE MULTI-TENANCY APPROACHES

There are several factors that influence the decision to adopt Multi-tenant database. These factors also help in determining the most suitable and appropriate approach of Multi-tenants database. The use of the system should be one of the influencing factors towards the decision. What is the area of usage of the system? Elmore et al (2011 p5) emphasise that the tenant application and usage requirements should be the primary consideration in deciding the right model of multi-tenant database. Sometimes users (tenants) are not equipped with necessary information about this before taking decision on what approach to adopt. Their decision is sometimes influence by what vendors tells them. There is need to examine all these basic factors before approaching a service provider in order to make the right decision on this.

Some of these factors are itemized by Keemti (2010) as follows.

1. Size of tenant database.
2. Number of tenants.
3. Number of users per tenant.
4. Growth rate of tenants.
5. Growth rate of tenant database.
6. Security.

7. Cost.

This research will examine each of the above factors as relevant and other factors identified by Yaish et al (2013) as follows:

1. Flexibility – ability to create multiple tables by tenants.
2. Time- time to build and configure.
3. Regulatory consideration (UK/EU countries)

According to Khan et al (2012 p41) while implementing multitenancy some factors must

be analyzed to successfully implement the multitenant applications which are:

1. System Cost
2. System Complexity
3. System Performance
4. System Security
5. System Scalability
6. System Flexibility

Some of these factors were actually mentioned by different scholars depending on the context of their research. It is possible to group them into four major headings as follows:

1. Economic (Time and Cost) consideration.
2. Growth consideration.
3. Security consideration.
4. Regulations consideration.

2.6.1 Economic Considerations

The economy of any project has to do with the cost and the time it takes to complete the project. Cost and time are major factors to be considered under this heading. Cost is vital when considering the appropriate approach to be adopted in the implementation of database multi-tenancy. This cost is referred to as total cost of ownership (TCO) which is broken down into three major types. These are infrastructural cost, management cost and application development cost (Wang et al 2008 pp94-95). Infrastructural cost includes the cost of hardware, software and utilization costs. And management cost are cost related to the operational activities and processes like lifecycle management, monitoring data backup and

restore while application development cost are cost related to meeting each customer additional unique requirement (Wang et al 2008 pp94-95).

Each of the three implementation approaches has different TCO depending on infrastructures, management and applications that are involved. Pippal (2011 p213) explained that separate database approach requires a high cost of maintenance since each tenant has separate database and each has some metadata used to relate to the correct database for each tenant. Chong et al (2006) illustrate economic consideration using two approaches: shared approach and isolated approach. The cost of shared approach tends to require a larger development effort because of the relative complexity of developing a shared architecture hence very high initial cost but in the long run the operational cost will be very low because they can support more tenants per server.

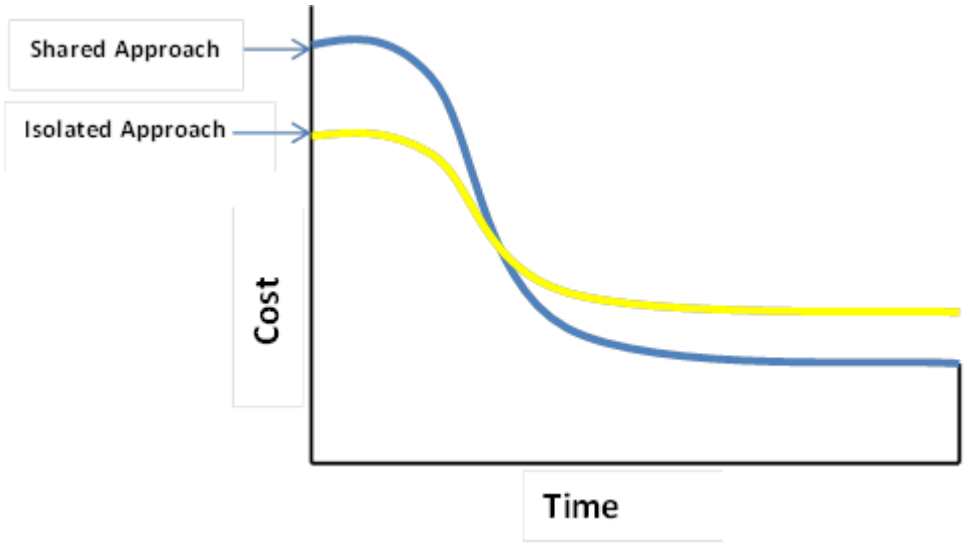


Figure 2.4 Cost over time for a hypothetical pair of SaaS applications; one uses a more isolated approach, while the other uses a more shared approach (**Chong et al 2006).**

Economic factors can constrain the development attempt which obviously can influence the choice of appropriate approach. And considering the time required to meet up with your customer (tenants) need, you might want to consider an approach that will require less time of development than a large-scale development approach. Aswathanarayana (2011 p4) also explain that cost saving is one of the major reason for going into multi-tenancy in SaaS. And when talking about economical consideration we will always illustrate it in terms of cost saving or ROI (Return on investment) and time. Aswathanarayana (2011) argue that every application instance usually consumes a certain amount of overhead memory and processing power, which amount to a significant cost with many customers. This is reducing as many customers (tenants) share a single instance. Another aspect of cost benefit is the cost

associated with each tenant in terms of licensing cost, in multi-tenancy everything can be run on a single software instance. All tenants are now run on one single software and it means only one software license will be required.

The advantage of several company or tenant with each varying number of users sharing the same set of IT resources both hardware and software in multi-tenancy database, this is fully utilized thereby reducing the capital expenditure that each tenant would have incurred in the business. Aswathanarayana (2011) support this by saying that multi-tenant environment maximizes the sharing of resources by allowing load balancing among tenants and reducing the CAPEX. There is efficient use of these resources because of the load balancing the idle time is completely remove or reduced to a very barest minimum.

The cost of managing the software and hardware is greatly reduced in multi-tenancy since there is only one instance of the application is involved. Instead of caring out upgrade of software and hardware for each tenants involved. Only one off upgrade maintenance is done. Aswathanarayana (2011 p4) puts it by saying multi-tenancy greatly reduces operational complexity and cost in managing the software by simplifying the release management process since there is only one instance of the application. Same upgrades are common for all customers and the package typically need to be installed only on a single server. All forms of management and maintenance are done once in a single location. Gao et al (2011 p324) illustrate the reason behind the widespread of multi-tenancy, one of which is the low cost for companies(tenants) in terms of hardware, software and utility of hosting centre (bandwidth, power, space etc.) and also lower cost of human resources to maintain processes and lifecycle via optimization and automation. The economic consideration in terms of cost is one factor so common in many scholars' prepositions. Aulbach et al (2011 p99) put it that the total cost of ownership (TCO) is reduced relative to the on premise solutions. The operational expenditure is reduced based on the fact that fewer processes are required for management. This invariably leads to reduction in capital expenditures and as well there will be increase in the resource utilization in the long run.

Bezemer et al (2010 p1) explained this economic factor by saying that multi-tenant model has twofold benefit; application development to the users becomes easier for the service provider and also the utilization rate of the hardware can be improved as multiple tenants share same resource. Bezemer et al (2010) also illustrate that these two factors invariably reduce the overall cost of the application and makes multi-tenancy concept interesting for customers

who might have limited financial resource. The concept is more financial benefits to the SME businesses who might not have financial weight to have on-premises solution or system for their operation, thereby given them the opportunity to get same resources from a service provider on the basis of pay per use.

Myer (2007) in Schaffner et al (2012 p157) brought more light into this discussion with illustration on cost of hosting one instance EC2 on Amazon's Elastic Compute Cloud with MySQL database for one year. This shows that it is very high for a single tenant system to bear. This is expressly shown with the diagram below. This is an indication that the more consolidation the better in terms of cost. This another evidence to support the fact that the more tenants you have on a system the better it is for the tenants as this will drastically reduce the subscription expenses. The cost itemised here is excluding the cost involved in administering the database.

Table 2.3-Yearly Cost to Host RightNow Database on Amazon EC2 (Schaffner et al 2012 p157)

	Single Tenant	Multi-tenant
Standard on-demand pricing	$2,233,800.00	$148,920.00
One-year reserved pricing	$1,470,900.00	$98,060.00

2.6.2 Growth Considerations

In multi-tenant database growth is very vital factor that must be considered when choosing one of the implementation approaches you desire. The nature, number and need of the intended tenants expected to render services to will definitely affect the type of multi-tenant database model or architecture you will eventually adopt.

I have decided to group all points 1-5 of the above mentioned factors as part of growth consideration factors. Aulbach and Jacobs (2007pp3-4) carried out experiment on memory (storage) and disk usage of the five different databases. This experiment focused on the shared process approach. It shows that memory storage required to handle 10,000 schema instances vary from 79MB to 2,061MB across the five databases.

Table 2.4 - Figure Storage requirements for schemas instances in megabyte (Aulbach and Jacobs 2007 p3)

	Memory 1 instance	Memory 10,000 instance	Disk 1 instance	Disk 10,000 instance
PostgreSQL	55	79	4	4,488
MaxDB	80	80	3	1,168
Commercial1	171	616	200	414,210
Commercial2	74	2,061	3	693
Commercial3	273	359	1	13,630

This shows that the size of tenant database in terms of storage capacity is a factor needed to be considered during the decision about choice of approach. How much storage space do you expect the average tenant's data to take? Schiller et al (2011 p118) give an illustration that the shared table approach has a promising for a service provider that has target for long tail because it offers the lowest over-head per tenant and thus suitable for a large number of small tenants. An example of 1000 tenants with each uses less than 50MB of data and at most 5 concurrent users. This is an indication that the number of tenants on the database and the number of users per tenant are all factors which must be thoroughly examined and also contribute to the performance of database system which invariably also a factor to be considered when taking decision on what approach to be used.

The degree of tenancy is an influence to the approach to be adopted. How many prospective tenants is your system going to serve? The rate of growth of tenancy ie at what rate the tenant will increase? Though you might not be able to estimate the use with authority assign to them but you must be able to reason out in terms of the size. Is it going to be a database for hundreds of tenants, for thousand, ten thousand or more? Myer (2007) in Schaffner (2013 p1) reveal that already in October 2007, the SaaS CRM vendor RightNow had 3,000 tenants which are distributed across 200 MySQL database instances with 1-100 tenants per instance. This means that the CRM can evolve overtime. It has been designed to accommodate more tenants. The growth rate of tenants on it is never limited to a small number. Lang et al (2012 p702) buttress this point by referring to multi-tenant setting that the degree of multi-tenancy becomes an additional factor that impacts performance that is experienced by every tenant on it. Lang et al (2012) says that per-tenant performance decreases with an increase in the degree of multi-tenancy but other way reduces the overall operating cost for the service provider. Two approaches were used by Lang et al to explain this which indicate the importance of

growth. Chong et al (2006) explains with a simple diagram where the choice will tend to consider these factors. The larger you expect your tenant base to be the more likely you will want to consider a more shared approach (Chong et al 2006). When all or some the of your tenant want to store very large amount of data, it is better to consider the isolated or separate-database approach. And isolated approach is also good when there are a larger number of concurrent users (end users) per tenant (Chong et al 2006). The diagram illustrates with arrows pointing to the direction the decision is likely to go.

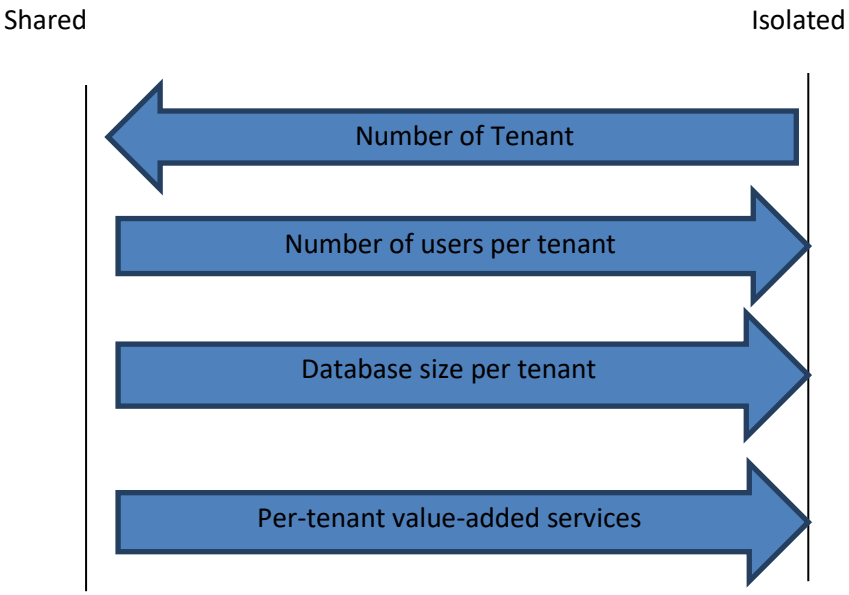


Figure 2.5 - Tenant-related factors and how they affect "isolated versus shared" data architecture decisions (Chong et al 2006).

Tenancy growth was also mentioned by Aulbach et al (2008 p1196) that a tenant denotes an organisation with multiple users, up to 10 for a small to mid-sized business. A simple web application like business email, a single blade server will support up to 10,000 tenants. This shows that the magnitude of growth is high in terms of tenant number. Therefore, if there are 10 concurrent users in each tenant, this will amount to 100,000 users in the database using the system simultaneously. Chong et al mention per-tenant value-added services. These services include backup, restore, customization reconfiguration etc. which are services that are easier to provide in the isolated approach. These services can still be provided in shared environment but not as easy as in the isolated approach. In mid-sized enterprise application like CRM, a blade server can support 100 tenants while a large cluster database can support up to 10,000 tenants (Aulbach et al 2008 p1196).

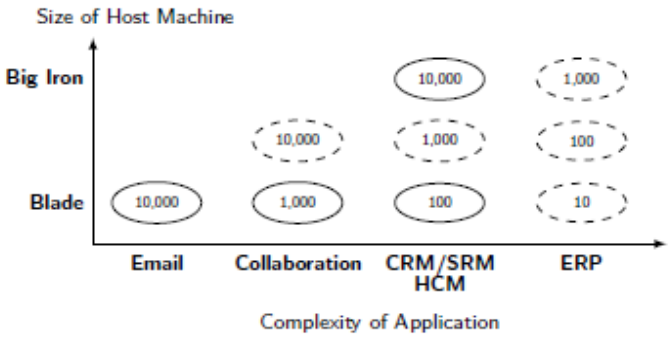


Figure 2.6 - Number of Tenants per database (solid circles denote existing applications, dashed circles denote estimate) source (Aulbach et al 2008 p1196).

Most multi-tenant database systems have challenges in coping with the increase in number of tenants using the same logical schema. Zhou et al (2011 p335) came up with DB2MMT system which provide a massive multi-tenant (MMT) database platform and provide case for multi-tenant application for long tail of tenants i.e very large number of small tenants using the same logical database schema. This also caters for consolidation of application or database for long tail of application with different schemas. This feature is an advantage for service providers who want to have very large number of tenant that would not require much memory space.

2.6.3 Security Considerations

In multi-tenant database system, one of the major concerns is the risk of data being exposed to the third parties. And because every database is design to store sensitive data, the prospective tenant will have very high security expectations. Every service provider will always want to operate to much higher security standard but sometimes this might not be to a hundred per cent. Therefore the service level agreement (SLAs) will have to provide very strong data security guarantees.

Some of the security issues related to multi-tenancy database include data isolation among the tenants. Gao et al (2011 p324) mention some of the challenges facing the ISVs (Independent Service Vendor/Provider) in delivering their service these include the data security isolation among tenant, the different tenants having different SLA demand, customization requirement and effective database scale out mechanism as the number of tenant increases. Hui et al (2009 p832) itemizes the problems faced in providing database as a service which includes security, contention for shared resources and extensibility. Hui et al (2009) also mentioned scalability as one of the problem which was defined as the ability to

serve an increasing number of tenants without too much query performance degradation. In spite of the increase in the tenancy and query request, the system should still be able maintain its performance level. This tenancy growth should not impact the service availability of other tenants.

Flexibility of the system is also a major security issue. A Multi-tenant database must be able to serve hundreds and thousands of tenants through one instance. Aulbach et al (2009 p881) expands this by saying that a multi-tenant database must be flexible by been able to extend the base schema to support multiple specialised version of application and to dynamically evolve the base schema and its extension while the database is on line. This means that the evolution of one tenant-owned extension should be independent on the service provider. The tenant should be able to achieve this without the service provider. The creation of multiple tables by individual tenant must be guaranteed. This is not most times possible because each tenant has its own system requirement such as objects attribute business logic etc (Gao et al 2009 p325).

Guo et al (2007 p2) also highlights two major challenges of native multi-tenancy design as firstly the system should support application level isolation among different tenants. This is quite in line with other scholars' positions. The multi-tenancy database should support more tenants and efforts are made to prevent the quality of service of one tenant from being affected by other tenants. While the processes and operation request of one tenant increases should not be a deterrent to other tenant on the platform. Secondly, Guo et al (2007) bring to notice that in this type of environment, the customization for one tenant should not impact on the other tenant during the runtime. This was also pointed by Vashistha and Ahmed (2012 p49) that Isolation should be carefully considered in almost all part of architecture design from both functional and non-functional level such as security, performance, availability, administration and also support tenant customization of their own services in runtime without impacting on others. Customization effort for one tenant which might require code modification and applications re-deployment, this sometimes affects other tenant. Once this becomes frequent occurrence will bring about the issue called service availability issues (Guo et al 2007 p2). The ISP will definitely have scores to settle with the customers (tenants) and will automatically results in termination of contract and withdrawer of service from such ISV.

Gao et al (2011 p325) gave a major requirement of multi-tenant database architecture as a system which is able to scale very large to support very large customer volume. This also Gao et al called incremental scalability, scale-out without impacting the service availability of other tenants. This particular feature is very difficult to achieve in most model types. This is also subject to the level of isolation in the approach adopted. Some service providers of multi-tenant database will have to put a lot of expertise and resources into making sure its infrastructure is as secure as possible against any form of data exposure, which would be very bad for its reputation.

One other security issue with this concept is that when one tenant has a particular application, it should be possible to deploy this application to other tenant in such system. This is a challenge mentioned by Zhou et al (2011 p335) as the cloud developer can easily deploy the existing application on to multi-tenant database without a large code change. When there is much code changes it affect other tenants' processes and subsequently lead to system down time.

Pippal et al (2011 p213) list out some concerns about third party offering multi-tenant service. One of which is the provision of a secure framework for authentication and authorization within the system. Pippal et al developed a protocol called Kerberos protocol. This incorporates a standard framework to implement authentication and authorization in the use of a multi-tenant database. This confirms who can access the database and verifies what you are allowed to do. This is to ensure what type of operation you are permitted to carry on the system. Authorization occurs after a successful authentication process. Every connection, login, query and update on a database must go through these two major security processes. These two security measures help find out why connection attempts are either accepted or denied. Authentication is important part of identity management while authorization has to do with access management. These are major mechanisms that must be enforced in a multi-tenant database for security purposes.

2.6.4 Regulatory Considerations

The database of every organisation is very vital to the growth and progress of the organisation. Enormous time and resources are devoted to the safety and protection of databases even as information is very important to organisations' processes. There are laws and regulations put in place by different governments that serve as protection to databases of different entities that operate in that geographical location. Companies, organisation and

governments are often subject to regulatory laws that affect their security and even record storage needs. The knowledge of these different laws and regulations is also to be considered when taking the decision about multi-tenant database. Chong et al (2006) argue that the investigation of regulatory environments that your prospective customers occupy in the market in which you expect to operate is important. This shows that you have to be conversant with the laws that operate in that area or country. It is important to find out whether there are any aspect of the law that present any condition that will influence your decision toward given your database service to a third party.

In 1996, the European Union (EU) finally adopted the EU Database Directive (Directive). The Directive created a two-tier protection scheme for electronic and non-electronic databases. Member states are required to protect databases by copyright as intellectual creations, or to provide a novel *sui generis* right to prevent unauthorized extraction or re-utilization of the contents of a database (Grosheide 2002 p39). The *sui generis* database right refer to as an intellectual property right that provides protection for the contents of a database (Schellekens 2011 p620). The difference between the two is that copyright infringement implies copying the structure, while the *sui generis* right infringement implies copying the contents themselves, irrespective of their copyrightability (Grosheide 2002 p39). This shows that provision is made for the protection of the content and also the structure of a particular database. The most important aspect of this law is the sui generis law which protect the content of the database.

Davison (2003 p10) explains these provisions of database protection with three different models. These models are; 1) Copyright protection is provided at a low level of originality, this means that a database user cannot take a substantial amount of the data contained within the database (Davison 2003 p10). There are prohibitions supported by law and regulation to those who might want to copy part of a database. 2) Copyright protection is provided if there is some creativity in the selection or arrangement of the database material which is back up with *sui generis* right (Davison 2003 p10). This is to say that prohibitions are placed on unauthorised extraction or re-utilisation of a substantial part of the data. This actually confers the exclusive property right to the database owner. 3) Copyright protection is provided for the creativity in the selection or arrangement of the database material (Davison 2003 p10). These models help in the protection of database in terms of security. Even when an unauthorised

someone gained access, these protections put in place by EU helps protect the database to some extent.

There are challenges when tenants are from different regulatory authority. Most of the countries have different laws which sometimes are not having same conditions and to now manage these conflicts will pose a great technical and economical challenge to the ISVs. One of the reason given by Johnson and Grandison (2007 p256) for privacy breaches and identity theft is the fact that there is weak and ineffective enforcement of the data protection laws as well as discrepancies and conflicts in the legal protections

2.6.4.1 Service Provider Restrictions

There are certain U.S regulatory frameworks that require data owners to ensure that their third party service providers are capable maintaining the privacy and security of personal information entrusted to them. According to Sotto et al (2010) there are two federal privacy laws that restrict the activities of service providers: they are Health Insurance Portability and Accountability Act 1996 and Gramm – Leach- Bliley Act. There are other state laws and regulations that impose security-based restrictions on service providers that have access to personal information. While these requirements do not restrict the geographic movement of a company's personal information (unlike the laws in the European Union), they do place restrictions on the use of service providers regardless of where they, or the data, are located (Sotto et al 2010).

2.6.4.1.1 HIPAA Restrictions on Health Data

Through its Privacy and Security Rules, HIPAA imposes significant restrictions on the disclosure of protected health information. With respect to disclosures to service providers, the regulations require covered entities to enter into business associate agreements containing statutorily mandated language before PHI may be disclosed to a business associate. Accordingly, any HIPAA-covered entity would first have to negotiate and enter into a business associate agreement with a service provider before it could store records containing PHI in a cloud computing facility. In some cases, HIPAA's substantive requirements could conflict with the cloud provider's terms of service, and a covered entity would risk a HIPAA violation by using such a provider for data storage (Sotto et al 2010). This is one of the America regulations used in protecting both parties involved in hosting of data.

2.6.4.1.2 Gramm-Leach-Bliley Act

Sotto et al (2010) alludes to the fact that for entities subject to GLB, the use of a cloud provider would be subject to similar restrictions. GLB's Privacy and Safeguards Rules restrict financial institutions from disclosing consumers' nonpublic personal information to non-affiliated third parties. Any such disclosures that are permitted under GLB are subject to a host of restrictions under both the Privacy Rule and Safeguards Rule. Pursuant to the Privacy Rule, prior to disclosing consumer personal information to a service provider, a financial institution must enter into a contract with the service provider prohibiting the service provider from disclosing or using the information other than to carry out the purposes for which the information was disclosed. Under the Safeguards Rule, prior to allowing a service provider access to customer personal information, the financial institution must (1) take reasonable steps to ensure that the service provider is capable of maintaining appropriate safeguards (i.e, the entity must undertake appropriate due diligence with respect to the service provider's data security practices); and (2) require the service provider by contract to implement and maintain such safeguards (Sotto et al 2010).

2.6.4.2 State Information Security Laws

A number of states impose a general information security standard on businesses that maintain personal information. For example, California requires businesses that disclose personal information to non-affiliated third parties to include contractual obligations that those entities maintain reasonable security procedures. Accordingly, covered businesses subject to the California law must contractually require cloud providers to implement appropriate safeguards (Sotto et al 2010). According to Sotto et al (2010) in 2008, Massachusetts issued regulations effective March 1, 2010 requiring any person who holds personal information about Massachusetts residents to develop and implement a comprehensive written information security program to protect the data. The regulations impose stringent and comprehensive data security standards on all businesses with Massachusetts consumers or employees. Companies are required to oversee service providers by (1) taking reasonable steps in the selection process to retain providers that are "capable of maintaining appropriate security measures to protect … personal information consistent with [the] regulations and any applicable federal regulations" and (2) contractually requiring service providers to implement and maintain appropriate security measures for personal

information (Sotto et al 2010). Companies subject to the Massachusetts regulations that are considering implementing a third party host solution must determine whether the service provider maintains appropriate security measures to protect the data to be stored and verify that the service provider's practices would not violate the company's own policies with regard to personal information.

2.6.4.3 State Breach Notification Laws

There are some laws and regulations that mandate service provider to notify tenants about any form of breach on its personal data. According to Sotto et al (2010) over 45 U.S. states and other jurisdictions have data security breach notification laws that require data owners to notify individuals whose computerized personal information has been subject to unauthorized access or acquisition. Having your data on third party system, the data owner may have little or no control over the security of company data being maintained, and it is impossible from for the data owner to confirm security conditions at the server location. Also, it is unclear how, or if, a data owner would be notified by the database service provider that its data had been subject to unauthorized access or acquisition that could trigger a notification requirement. So, these laws are put in place to make service provider more responsible for tenant's data.

2.6.4.4 Breach Provisions under HITECH Act

The Health Information Technology for Economic and Clinical Health Act, established new information security breach notification requirements that apply to a wide range of businesses that handle PHI and other health data. The regulations apply to all breaches discovered by covered entities and business associates, but include a harm threshold limiting the breach notification requirement to breaches that present a significant risk of harm (Weis and Alves-Foss 2011). This was further explained by Sotto et al (2010) that any third party service providers (including cloud providers) must notify covered entities to which they provide services of any breaches they discover so the covered entity can comply with the notification requirements. To the extent a HIPAA covered entity discloses PHI to a service provider, it risks exposure to federal data security breach notification requirements provided under the HITECH Act.

2.6.4.5 European Union Regulations

The European Union data protection authorities have put some regulations in place to ensure data protection compliance by both vendors and prospective users of services of the providers.

In the European Union, we have either a data controller or a data processor. Data controllers determine the purposes and means of the processing of personal data and are responsible for compliance with data protection law, whereas data processors process personal data on behalf of controllers (Sotto et al 2010). In terms of cloud computing, characterization of an entity as a controller or a processor may depend on the type of cloud computing system that is used or on the technical setup of the system (Sotto et al 2010). This determines the liability of the respective parties for compliance with data protection obligations. But more significantly, the controller be responsible for discharging data protection obligations even in the case where the data has been outsourced or transferred to a third party—including a cloud vendor—for processing. Companies are therefore expected to undertake a rigorous assessment of its responsibility for the personal data processed by the cloud provider. And once this is ascertained then they are met to go into a data processing agreement requiring the service provider to act only according to the company's directives in order to make sure that there is an adequate technical and organisational security and otherwise to comply with legal requirements (Sotto et al 2010).

2.6.4.6 Data Transfer Restriction Issues

There are some jurisdictional issues in the context of cloud computing or third party service provider. The EU has restrictions place on the international transfer of personal data by EU member states. Transfers of personal data outside of the European Economic Area that originate within the EEA are prohibited unless the receiving country provides for an "adequate" level of protection (Sotto et al 2010). Currently, the European Commission considers only a handful of countries to provide an adequate level of data protection, and the United States is not one of them. The transfer of personal data to a country that is not considered "adequate" may be authorized if the data recipient has implemented a legal mechanism providing for an adequate level of protection (such as adherence to the U.S. Safe Harbor Program) or if the data controller can rely on an exception to the prohibition (Weis and Alves-Foss 2011). Such mechanisms are challenging to implement in a cloud context and

may require the approval of an EU data protection authority. To seek to address these concerns, some cloud vendors offer segregated EU clouds that keep EU personal data from being transferred outside the European Union (Sotto et al 2010).

2.6.4.7 Information Security Safeguards

According to Sotto et al (2010), EU data protection law requires data controllers to implement appropriate technical and organisational measures to protect personal data against

(1) Accidental or unlawful destruction or loss;

(2) Unauthorized alteration, disclosure or access (in particular where the processing involves the transmission of data over a network); and

(3) All other unlawful forms of processing.

It is important that in using any third party service like Database as a Service (DaaS), tenant (companies) should consider the possibility of potential unauthorised disclosure or access. Therefore, measure like authentication and access safeguards must be robust and make available for a high level security. According Sotto et al (2010) individuals have a fundamental right under European Union data protection law to access, block, rectify or delete their personal data. Due to the technical set-up of a cloud computing infrastructure, it may be difficult to guarantee that requests for access, blocking, rectification, or deletion are effectively and properly managed. A service provider agreement would have to address this issue specifically. These are part of what the SLA address which is made very clear and agreed upon at the beginning of the contractual agreement between the tenant and service provider.

2.7 Summary

This Chapter has presented a review of the literatures in relation to the concept of MTD. The overview of the concept meaning, the requirements for its implementation, the challenges associated with it, the migration of database in MTD environment, the implementation of MTD in cloud environment were all examined in this chapter. The different implementation models of MTD available were examined in respect to their features and comparison was also made between the different models. This Chapter also present other technologies where comparison were made with the MTD and looked into what the literatures have said about the

evaluation and adoption of these technologies. This Chapter also looked into other technologies similar to MTD and examining the issues surrounding their adoption.

There is need to identify the factors that influence the adoption of MTD in order to achieve a leading step to the aim and objectives of this research. This Chapter has also presented the various factors based on exploration of available literatures in the subject area. These factors are further grouped into four for the purpose of this research which include economic, growth, security and regulation factors. Finally, each of these groups was extensively reviewed in this Chapter. Some regulations were also presented to illustrate the importance of regulations in different part of the world.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 INTRODUCTION

Saunders et al (2009) describe research as "*something that people undertake in order to find out things in a systematic way thereby increasing their knowledge*". They consequently identified the characteristics of a good research which include: ensuring data are collected systematically, data are interpreted systematically and there is a clear purpose to find things out. There are different styles of research such as constructive, theoretical, empirical, nomothetic, idiographic, critical, information systems research and so on.

The research methodology is the various ways, methods, designs and systems which researchers use in administering and collecting data in a research study (Ezejelue and Ogwu, 1990). It is therefore important to outline the methodology adopted to achieve the above stated characteristics of a good research. This Chapter gives an overview of the approaches and the methods adopted for the capture of data needed, steps of interpretation and clear systematic steps to achieve the aim and objectives of the research.

3.1 Research Approaches

When conducting a research, it is necessary to determine which approach is being implemented. Therefore research approach is the involvement of either the steps to the development of a theory or the statement of parts of an existing theory in testable form (Singh and Bajpai 2008 p11). This research employs both inductive and deductive approaches of research. The inductive part involves the observation on the Multi-tenant database literature reviews which form a pattern, a tentative hypothesis were formed. The combination of these hypotheses forms a theory in the form of a framework. While the second part is deductive approach which from the framework the hypotheses were observed in the form of survey using questionnaire. This observation were analysed and the results will lead to the confirmation of the framework. The framework will not only be confirmed but also adjustment will be made to it to correspond to the findings from the analysis.

3.2 Research Method

Methodology is the study of how research is done, the principles that guide our research practices. McGregor and Murname (2010 p2) says "The word methodology comprises two nouns: method and ology, which means a branch of knowledge; hence, methodology is a branch of knowledge that deals with the general principles or axioms of the generation of new knowledge. It refers to the rationale and the philosophical assumptions that underlie any natural, social or human science study, whether articulated or not. Simply put, methodology refers to how each of logic, reality, values and what counts as knowledge inform research."

Data was collected from different sources using various methodologies. According to Blaxter et al (1996), the data required can be classified as qualitative if it comes in word form, while they are regarded as quantitative if they come in the form of numbers. This was also explained by Saunders et al (2007) that quantitative data refers to every numerical data or contained data that will be helpfully quantified in helping the researcher answer research questions and meet the objectives and can be a product of all research strategies such as experimentation, examination, case study, action research, stuck theories, ethnography and archival research. Qualitative data is each and every one of the non-numeric data or data that cannot be quantified and can be a product of all research strategies such as experimentation, investigation, case study, action research, stuck theory, ethnography and archival research (Saunders et al, 2007).

However, this research employs a mixture of both qualitative and quantitative research. This involves several issues relating to the concept of Multi-tenant database which are broadly reviewed from several scholar articles. These are not numerically quantifiable and moreover, this process involves experts and their organisations. These are tangible issues such as experience of the experts and operations of the organisations which play a vital role. Therefore, a combination of both strategies are used here to cover statistics (quantifiable data) and experts' knowledge and experiences (abstract data) in order to satisfactorily achieve the aim and objectives and also complete the research.

3.2.1 Data Collection Method

This research involves multiple research questions and the researcher has applied diversity in the source and types of required data. Therefore, as in many other research areas, the

researcher uses both primary and secondary sources of data. According to Cameron (1999), from the primary data, the researcher collects either by direct observation, measurement, interviews, questionnaires or other means which can be modified to his requirement to give answers to exactly the question which concern him, from a suitable sample while the secondary data is a process of reanalysing data that have previously been collected for some intention (Saunders et al, 2009). They are other people's facts and figures, which may be surveys, carried out by other people; sets of government information such as population census, company report, academic research journal report, etc. The utilization of secondary data saves time and money, and can be beneficial because part of the background needed for the research has been already surveyed with a pre-established degree of validity and reliability. The researcher can re-use them without the need for re-examining them. However, using the secondary data may not be quite adequate for the research questions because they have been collected for other studies with diverse objectives (Craig and Douglas, 2000) but it can provide a baseline for a research which is about to start and be useful in designing the appropriate methodology by identifying key issues and data collection methods.

This is why a critical review of relevant literatures was done as one of the sources for collecting the secondary data. In this research these sources included books, journal articles, online data sources, document and catalogues of MTD services provider with multi-tenant database features and products. A selected pilot sample group was used to collect primary data in order to contextualize and strengthen the arguments in the research by providing both quantitative and qualitative evidences. This also helped the researcher to analyze the collected data from primary sources about this particular environment more robust and in-depth.

3.2.2 Chosen Data Collection Methods

The selected research method for this research follows a mixed methods approach which involves five main stages as follows:

1. A review of the literature review to establish the knowledge gap.
2. The development of a questionnaire as an outcome of the literature review and the distribution of this questionnaire to a pilot sample group.
3. The development of a framework from the literature review.

4. The development of a questionnaire as a results of the framework and administer online to sample size that expand to all continents of the world.
5. The development of sets of questionnaire administer online to experts in the field including database administrators within MTD user organisation to validate the research findings, recommendations, guidelines, the framework and the expert system.
6. The exploration of interview with database administrator within MTD user organisation to validate the research findings, recommendations, guidelines, the framework and the expert system.

The reason behind the choice for a quantitative research strategy that involves the use of a survey is to answer most of the 'what and who' questions that are associated with database multi-tenancy implementation and adoption. Then, there will be an in-depth qualitative investigation of issues informed by the survey in answering most of the 'how' and 'why' questions.

One major reason for using survey is that surveys are viewed as the most appropriate method of studying participants' behaviour and job perceptions (Mintzberg, 1973; Rea and Parker, 1997). Also Rea and Parker (1997) said there is no better method of collecting information about a large population other than survey.

According to Rea and Parker (2014), a good questionnaire is designed by considering how it will meet the research objectives, to obtain the most complete and accurate information possible, to make it easy for respondents to give the necessary information, to arrange the question in order to have sound analysis and interpretation possible, and finally to keep the questionnaire brief, concise and straight to the point. These are factors taken into consideration while developing the questionnaire and more importantly the supervisory team were all involved to guide the researcher in the development of the questionnaire. This makes the design to be properly done in accordance to research standard ethics.

Bryman (2008) indicate that survey is used to collect data when the information sought is reasonably specific and familiar to the respondent. Also Denscombe (2007) alluded to this point by saying that questionnaires can be used to provide information on a particular point of interest in a research for easy analysis and interpretation. Based on these, these set of data were collected because basic information was required from experts in this field of research to help in arriving at findings from the analysis and interpretation done on the data.

3.3 Literature Review

The importance of carrying out a concise literature review was highlighted by Gall *et al*. (1996); Fellows and Liu (2003).

i. Compare the research idea to existing knowledge in the chosen field;
ii. To provide insight to whether the research is viable;
iii. To provide insight to whether the research addresses a topical issue and will also not lead to repetition;
iv. To help redefine the research topic to reflect current trends in the chosen field;
v. To help develop an appropriate methodology for undertaking the research; identify and /or suggest routes with regard to advancing knowledge;
vi. Help generate other ideas;
vii. To assist in further refinement of the research questions and objectives.

This stage (literature review) helps in the complete review of various literature that concern database multi-tenant management, implementation and adoption. This provides the background and context upon which this research is set up; also the findings from the literature review will form a strong basis for the potential field survey.

3.4 DESIGN OF SURVEYS:

There are two types of survey used in this research which are questionnaires and interview.

3.4.1 Questionnaire

The questionnaire is a data collection technique whereby each individual is requested to respond to similar series of questions in a prearranged order (DeVaus, 2002 as cited in Saunders et al, 2009). The purpose is to enable the researcher recognize and explain the degree of disparity in answers on exact topics and to find for any relationships between views on one set of questions to position on another (Cameron, 1999). Questionnaire use the descriptive research such that it undertakes using approach and views questionnaires of organisational procedures which will facilitate the researcher in recognizing and illustrating the inconsistency in different phenomena (Saunders et al, 2009). The types of questionnaires is described in Figure 3.3 below

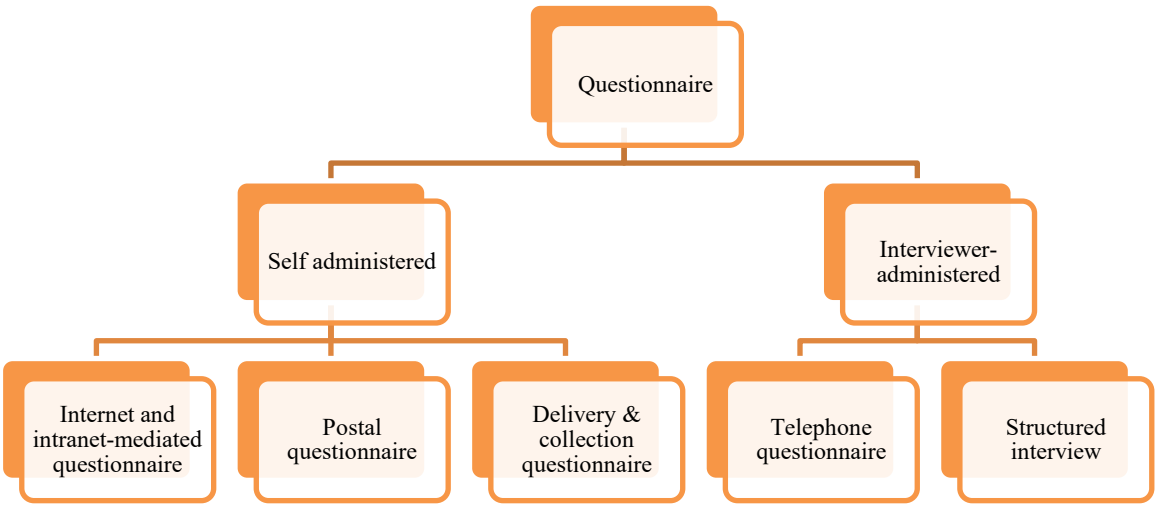


Figure 3.3 Types of Questionnaire (Source; Saunders et al 2009)

This was designed carefully to make sure that it demonstrates and draws out useful responses to the questions it's possess. Oppenheim (1992) and Baker (2003) advocate for best practice in designing a questionnaire, and such practices include that the questionnaire should be made to be easy to read and understand (set up in simple language), should be short and have the potency of being completed quickly, and also organized/arranged to flow easily without any form of bias. Accordingly, Knight and Ruddock (2009), when developing a questionnaire there should be careful consideration in its design. This is to ensure that the data generated can be analysed in the way the researcher wishes, which is mostly done through statistical approach. Furthermore, the questions included must be constructed to ensure reliability and validity of the information obtained, and is designed in a very simple form (Peterson, 2000). This demands that they should be brief, relevant to the topic, clear and unambiguous, specific, and objective (Peterson, 2000). In view of the feedback expected, both open-ended and closed-ended questions were applied. In other words, the questionnaire consist of questions that just require 'ticked-boxes' and questions whereby the respondents was required to administer their personal and independent inputs. Once formulated, a draft questionnaire should be pre-tested since this procedure is vital to its success as a research instrument. Initially, the pilot should be with another knowledgeable and academic person (the supervisor of a research project), and with professionals and experts in the field.

In this research, the questionnaires were systematically administered in United Kingdom, as a delivery and collection questionnaire under self-administer type to representatives from:

- UK Oracle users group:

Also, the questionnaires were administered online as internet and intranet-mediated mode of questionnaire under self-administer type to different set of respondent across all continent of the world which includes:

- Administrators from different organisations.
- Users from organisations.

3.4.2 Interviews

The interview will be given to some representatives of various organisations and groups as in the case of questionnaires. However, the interviews could run concurrently with the questionnaires; and will still continue after responses from the questionnaires have been received. The use of in-depth interview will present clearer answer which will help in the interpretation of the responses.

A 'face-to-face' pattern of interview will be adopted by the researcher because, according to Creswell (2003), it gives room for better interaction between the researcher and the interviewee, and also for better observations to be made. However, in a situation that the interviewee could not be contacted on a 'one-on-one' basis, the alternative method will be 'telephone interview'.

Oppenheim (1992) states that there are about four formats for structuring interviews and they include totally structured; semi structured; open questions with structured answers; totally unstructured. For this research, the "SEMI STRUCTURE" will be designed and utilised. It involves structured question with open responses, and gives the interviewee the opportunity to express himself/herself without bias.

3.5 Data Analysis:

Data analysis is the process of bringing order, structure and meaning to the mass of collected data (de Vos et al., 2002). Firstly, a structured literature review was conducted in order to provide the academic foundations of the research after which descriptive statistics were utilised to analyse quantitative data in the first and the second phase of the data collection. Taylor (2005) comments that descriptive statistics are used to describe quantitatively how particular characteristics are distributed among a group of people and that researchers use

descriptive statistics when reporting the findings of a study. Descriptive statistics are used to organise and present data in summary form (Taylor, 2005).

The Statistical Package for the Social Sciences (SPSS) was utilised particularly to carry out the descriptive statistics analysis.

In analysing the first phase of the data collected from the focus group, a simple quantitative statistical method was adopted known as weighted score method also known as numerical indicator (Abeysekera 2001 p10).

This research on MTD is largely quantitative and is concerned with measurement of majorly the nominal and ordinal variables. The data from the survey were coded into SPSS and represented in numerical values. These data were subjected to the following descriptive statistical tools;

1. Percentage Frequency Distribution.
2. Cross Tabulation.

A predictive analytical method called Relative Importance Index (RII) was also used on the second phase of data collection to also determine the degree of impact of each factor. More details about this are found in the chapter six.

3.6 Overview of the Entire Research Process

The entire research process follows a series of steps which systematically leads to the development of the scientific guideline for the drive towards multi-tenant database adoption. Figure 4.4 below presents the entire research process for this study showing the various steps involved in carrying out the research.

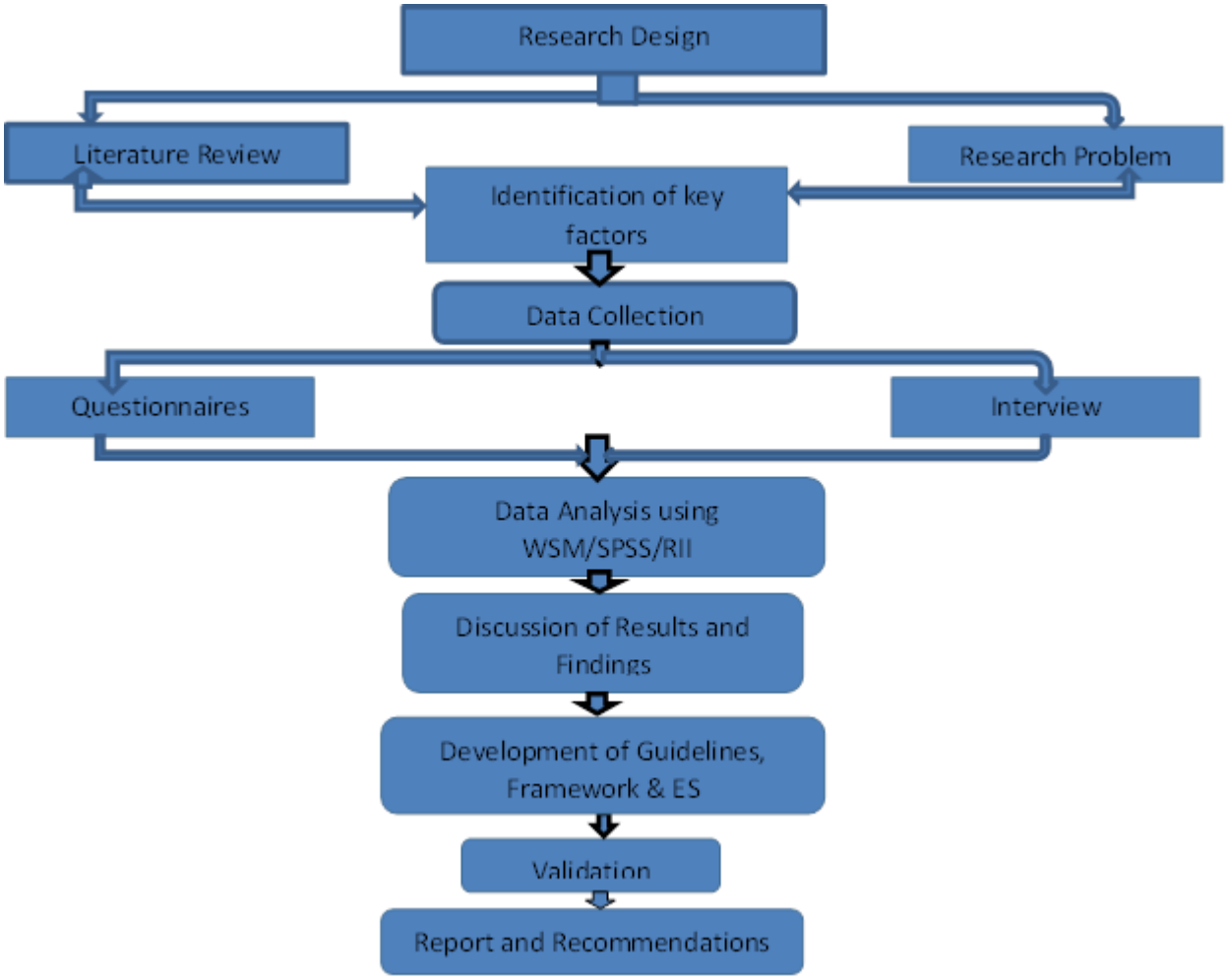


Figure 3.4 - Research process

3.7 Summary

This Chapter has outlined the methodology and research design employed in this research. The establishment of a robust research methodology was the primary focus of this Chapter. The research approaches were thoroughly explained and the reason for the choice made in this study was also explained. The research combined both quantitative and qualitative research methods with a greater dominance on the quantitative method. There was an in-depth explanation of the research method used for the purpose of this research and a justification for the choice of the quantitative research mode. Also, the Chapter has identified the various data sources that were used in the research. The choice of the research strategy was based on the nature of the data. A survey in the form of questionnaire was adopted in the first and second phase of the research. Also, ethical concerns with regard to data collection were emphasised as well as the method of data analysis, highlighting the various steps that were involved.

CHAPTER FOUR

4.0 The Conceptual Framework

What is a Conceptual Framework?

Every research work is based on a conceptual framework or model. This is the reason while this Chapter is dedicated to the framework of the study. The conceptual framework of a study is the system of concepts, assumptions, expectations, beliefs, and theories that supports and informs your research which is a key part of a design (Miles & Huberman, 1994; Robson, 2011). Miles and Huberman (1994) also defined a conceptual framework as a visual or written product, one that "explains, either graphically or in narrative form, the main things to be studied, the key factors, concepts, or variables and the presumed relationships among them". There are variable considers in this work which are the factors identified in the earlier Chapter.

Another school of thought explains framework as the actual ideas and beliefs that you hold about the phenomena studied, whether these are written down or not; this may also be called the "theoretical framework" or "idea context" for the study (Jabareen 2009 p50). According to Morse et al (2002 p1) these are characteristics identified from previous inquiry that provide an internal structure which has a starting point for observations and interview questions, and for analysis. The researcher proceeds by building on these structures or categories, padding them out or "giving them flesh" and organizing the ways they fit together. In this research, a graphical representation of ideas from different scholar is presented. This formed the basis of the investigation to validate or restructure the framework for benefit of the intending MTD subscribers. This is further explained by Jabareen (2009) that conceptual frameworks are products of qualitative processes of theorization.

The most important thing to understand about a chosen conceptual framework is that it is primarily a conception or model of what is planned to study, and of what is going on with these things and why, a tentative theory of the phenomena that is being investigating (Jabareen 2009 p50) . The function of this theory is to inform the rest of the research design, to help assess and refine the goals, develop realistic and relevant research questions, select appropriate methods, and identifies potential validity process to the conclusions. This thus helps justify the research.

4.1 Economic Factor and its Hypothesis

This aspect is the cumulative consideration of cost and time. The goal of every organisation is to find a cost-minimal solution for adoption. This is one of the advantages pointed out by Walraven et al (2014) that multi-tenancy achieved high operational cost efficiency by sharing the same resources among multiple customer organisations. In discussion of multi-tenancy in relational databases, Narasayya et al (2013) also attest to this point giving the example of Microsoft SQL Azure where resources of a single database server can be shared among multiple tenants to achieve cost reduction. Multi-tenant is crucial for this purpose since dedicating a machine for each tenant makes the service prohibitively expensive. Yaish and Goyal (2013 p939) also allude to the fact that multi-tenant database helps reduce the Total Cost Ownership (TCO) from the tenants' perspective, by sparing them from spending money and effort on writing SQL queries and backend data management codes. According to Khan et al (2012 p41) the primary factor which is highly influential on Multi-tenant applications in a positive way is cost. Instead of dedicating separate infrastructure for each individual tenant the infrastructures are shared by virtualisation. Khan et al (2012) said this result in significant decrease in system cost. In the study by Elmore et al (2011), it was pointed out that elasticity, pay-per-use, low upfront investment, low time to market, and transfer of risks are some of the enabling features that make MTD a ubiquitous paradigm for deploying novel applications which were not economically feasible in traditional enterprise infrastructure settings. This obviously will drive users towards the adoption of MTD. This has proved that cost is minimised in the adoption of a MTD which means that subscribers will want to embrace the concept for the purpose of saving cost.

Also considering the time it takes to deploy and configure an on-premise database system which will normally take much longer time and effort to achieve. Therefore, organisation will want to adopt a ready-made database system and save the time. And considering the time required to meet up with your customer (tenants) need as a database service provider, you might want to consider an approach that will require less time of development than a large-scale development approach. Das et al (2013) explains that in a multi-tenant environment, the tenants tend to use the application base on the fact that they enjoy benefit of economy service with resource demands becoming very quick with the cost benefit. Looking at time and cost as economic factor, based on all the position of different scholars, it shows that organisations (tenants) will choose MTD from any reliable vendor as a means of storing and protecting

their database. This means that economic factor will drive tenants towards the adoption of MTD.

Hypothesis 1 – Economic factor drives towards MTD adoption.

ECONOMIC FACTOR ⟹ MTD ADOPTION

4.2 Security Factor and its Hypothesis

Gao et al (2011) explain security in terms of isolation among tenants. Gao et al argue that access to other tenants' data is sometimes granted by the system admin but if not controlled can be a source of data leakage. There should be a more flexible, fine-granular access control mechanism which will prevent SQL injection among tenants. Furthermore Gao et al highlight that MTD has issues related to customisation of applications to the business needs of every tenants, considering the fact that each tenant might want different SLA (Service Level Agreement). The control of this data leakage and management of the SLAs of different tenants are issues related to the security of the MTD which could discourage its adoption.

Khan et al (2012) allude to the fact that security is a major concern when it comes to multi-tenancy. There are two approaches of security which are; the overall system security and the individual security level of a tenant (Khan et al 2012). Once one of the security approaches is affected, it will also affect other and the system as a whole. The system architecture should be designed such that it should be scalable considering the future requirement. If this is not done it will affect the security of the system. And security is a key credential when it comes to a standard system. But security is compromised by sharing of resources to achieve multi-tenancy (Khan et al 2012). So according to Khan et al resources sharing makes application more scalable and flexible but doesn't give one hundred percent secured system. According to Kshetri (2013), the primary concerns related to security, privacy and confidentiality might generate huge cost implication that might outweigh the benefits of this concept. This now indicate that the financial implication of security breaches might too much compare to the cost invested in an on-premises database.

Pippal and Kushwaha (2013 p8) proposed approach for multi-tenant database is the shared database shared schema technique which has the advantage of suitability for large number of tenants. Cost minimised and leveraged benefit of using same hardware, software, database, schema and table for all tenants. Pippal and Kushwaha (2013) proves that this guarantee good isolation and security. While the second approach shared database separate schema will be

limited to the number of instances of the database supported by the DB server. This now means that adding more tenants might result in security lapses which could very expensive to resolve.

However, the use of shared database shared schema technique requires the implementation of extension table. The extension table contains a lot of information for Meta data ie for a single row of table of a tenant that consists of four columns, the Record_ID and Extn_ID are repeated four times, and this information introduces a kind of redundancy (Pippal and Kushwaha 2013). Li et al (2014 p32) argue also that to guarantee high level of performance in terms of data integrity of results from query operations in a large multi-tenant database, adequate pivot table for tenant data are set up to speed up the query process. This will lead to having several stored tables which might not be in used after the query operations. This is data redundancy. Data redundancy is one of the major issues that affect organisation data management and security. When queries are issued depending on the number of queries the processing time is increased because of this factor.

Luo et al (2015) argues that a lot of memory is consumed on maintaining a huge number of tables for a large number of inactive tenants. This will unavoidably limit the number of tenants that can be consolidated into the same DBMS instance. This implies that as time goes on so many tables would have been created and while few of these will be used or queried. These idle tables would occupy a lot of system memory (Luo et al 2015). This will invariably lead to degraded system efficiency and scalability. As the number of tenant increases the number of table also increases and the system becomes very slow and inefficient. The performance of the system increasable becomes very slow, not meeting up with the demands of tenants. This is a major issue that lead to lack of security. This is also one of the reasons for the slow pace towards the multi-tenant database.

Luo et al (2015) identifies the issue of scalability which refers to the ability to serve an increasing number of tenants to get an acceptable level of performance is a major problem of multi-tenant database. A previous study by Aulbach et al (2008) also observed very high significant performance degradation on a database server when the number of table exceeds 50000. When more tables are created by more tenants, the response time to tenant's request will become increasingly slow as the number of tables increases.

Based on the points of the above scholars, Security issues will drive tenants towards rejection of MTD.

Hypothesis 2 – Security factor drives towards MTD rejection.

SECURITY FACTOR ⟹ MTD REJECTION

4.3 Growth Factor and its Hypothesis

Growth of Multi-tenant database is an aspect that DaaS providers always want to handle before it impedes the smooth running of their platform. The growth is in different aspects which include the increase in the number of tenant, the increase in the storage capacity of each tenant and also the increase in the number of users per tenant. These seem to be concerns to both the users (tenants) and service providers.

According to Narasayya et al (2013), there is a serious problem with varying and increasing tenant's workload contending with other tenants for key resources such as CPU, I/O and Memory. Tenants of a relational database-as-a-service platform can execute arbitrary SQL queries that can be complex and whose resource requirements can be substantial and widely varied. This tends to affect the performance of the MTD system. This is because the performance of individual tenant's workload can vary significantly depending on the workload of other tenants contesting for the shared database resources. Narasayya et al (2013) further illustrate that tenant's data size, distribution and access pattern can change over time. According to Nasarayya et al these factors contribute to even greater variability in query throughput and latency. When the throughput (ie queries/sec) is affected then performance is also affected and hence leads to a very slow system. This suggest that there must be a mechanism in place by MTD providers to reduce to barest minimum the variability in the performance which occurs as a result of contention with other tenant for very important shared database resources.

Lang et al (2012) refers to growth as the degree of multi-tenancy which is explain to mean the increase in the number of tenant and the increase in the storage space occupied by each tenant. Lang et al (2012 p702) alludes to the fact that the degree of multi-tenancy is an additional factor that impacts performance both for the overall system and the performance experienced by each individual tenant. The increase in the degree of multi-tenancy decreases per-tenant performance. However, there is a high reduction in the overall operating cost for the MTD providers. The MTD providers always adopt scheduling method of provisioning the resources to accommodate the tenants for each resource. This will help improve the performance level and thereby still enjoy the reduced overall operational cost for the MTD

providers which also imply that the tenants' cost will also be reduced. This provisioning of resources and other measures to minimized the variability in performance will encourage more tenants subscribe to the technology MTD.

Gao et al (2011p327) also explain along this line that the maximum number of tenants supported by multi-tenant system should increase in direct proportion to the increase of resources. This will ensure that the performance level of each tenant will remain in an acceptable level. This provides a more cost-effective and smooth scalability that will enhance system performance. In addition to what Aulbach et al (2008) said that MTD has a limited number tables it can handled depending on the amount of memory available, but this is ameliorated by introducing the shared table method. All these points based on the growth further encourage the adoption of MTD. Growth issues will drive tenants towards adoption of MTD.

Hypothesis 3 – Growth factor drives towards MTD adoption.

GROWTH FACTOR  MTD ADOPTION

4.4 Regulation Factor and its Postulate

There are laws and regulations put in place by different governments that serve as protection to databases of different entities that operate in that geographical location. Companies and organisations are often subject to regulatory laws that affect their security and even record storage needs. According to Bezemer and Zaidman (2010) individual country may have its own legislation in place, example is the European Union's (EU) legislation on the storage of electronic invoicing, which states that electronic invoices sent from within the EU must be stored within the EU as well. Therefore, the knowledge of these different laws and regulations are considered when making decision about multi-tenant database. The entire responsibility of database management, i.e., database backup, administration, restoration, database reorganisation to reclaim space or to restore preferable arrangement of data, migration from one database version to the next without impacting availability which befall organisation is now transferred to database as a service provider (Hacigümüs et al 2002). There is variability in terms of where the physical data resides, where processing takes place, and from where the data is accessed. Given this variability, different privacy rules and regulations may apply. Because of these varying rules and regulations, by definition politics

becomes an element in the adoption of Database as a Service, which is effectively multijurisdictional (Avram 2014). This aspect of administration, restoration and database reorganisation will require a protecting laws or regulations that will safe guard the interest of both the tenant and the service provider. There are federal, state, and local laws and regulation. And according to (Weis and Alves-Foss 2011) how each of these applies depends on the situation, the service-level agreement (SLA) between the customer and provider, and data location. The latter brings a new concern because users don't control and often don't know where their data is physically located. This introduces a slew of legal problems, especially if their data crosses international boundaries. This shows that the service- level agreement will take care of all the security flows that may occur as a result of regulations.

The SLAs for SaaS application differ per tenant. According to Gey et al (2015), the SLA violation and service disruption caused by updating and upgrading the entire SaaS application for the sake of maintenance and evolution is unacceptable. This cause service to be disrupted at the tenant's point during the run time and eventually results in multiple co-existing versions of individual components and such brings about what we might refer to as management complexity. However, Gey et al (2015) proposes a powerful run-time adaption mechanism that will allow the provider to update and upgrade the platform in a gradual process on a fine-grained, per-tenant basis. To maintain a good quality and also comply with SLAs throughout an upgrade, the multi-tenant SaaS application must ensure on-going tenant request and also keep accepting and processing new request. These are covered by the SLA agreement and understanding that exist between the service provider and the tenants.

Considering the complex regulatory issues surrounding data protection across various jurisdictions, the inability to know where one's data is located, or if and when the data may be moved to another state or country, implies a good deal of potential legal risk (Sotto et al 2010). There are numerous state and federal privacy and data security laws and regulations mentioned earlier on requirements that are triggered as result of storing your data on a third party platform. As mentioned earlier, the hosting of data owned by another tenant belonging to a different data protection authority will adequately be taken care when the host provider operates what is called Safe Habour Program e.g U.S Harbour Program (Sotto et al 2010). This means that the receiving country or state or authority where the provider resides is protected.

In conclusion, this shows that irrespective of the state or authority under which your service provider occupies, there will be a protective regulation to safeguard the tenant's data. However, where a safe habour program is absent there will be a discouragement to MTD. Prospective tenants would not want to take the risk of given her data to a third party host who cannot guarantee the safety of it. Therefore regulation should either encourage adoption or rejection of Multi-tenant database.

Hypothesis 4 – Regulation factor drives towards MTD adoption or rejection.

REGULATION FACTOR → MTD ADOPTION

REGULATION FACTOR → MTD REJECTION

All these four hypotheses are combined to form the MTD Conceptual Framework presented below.

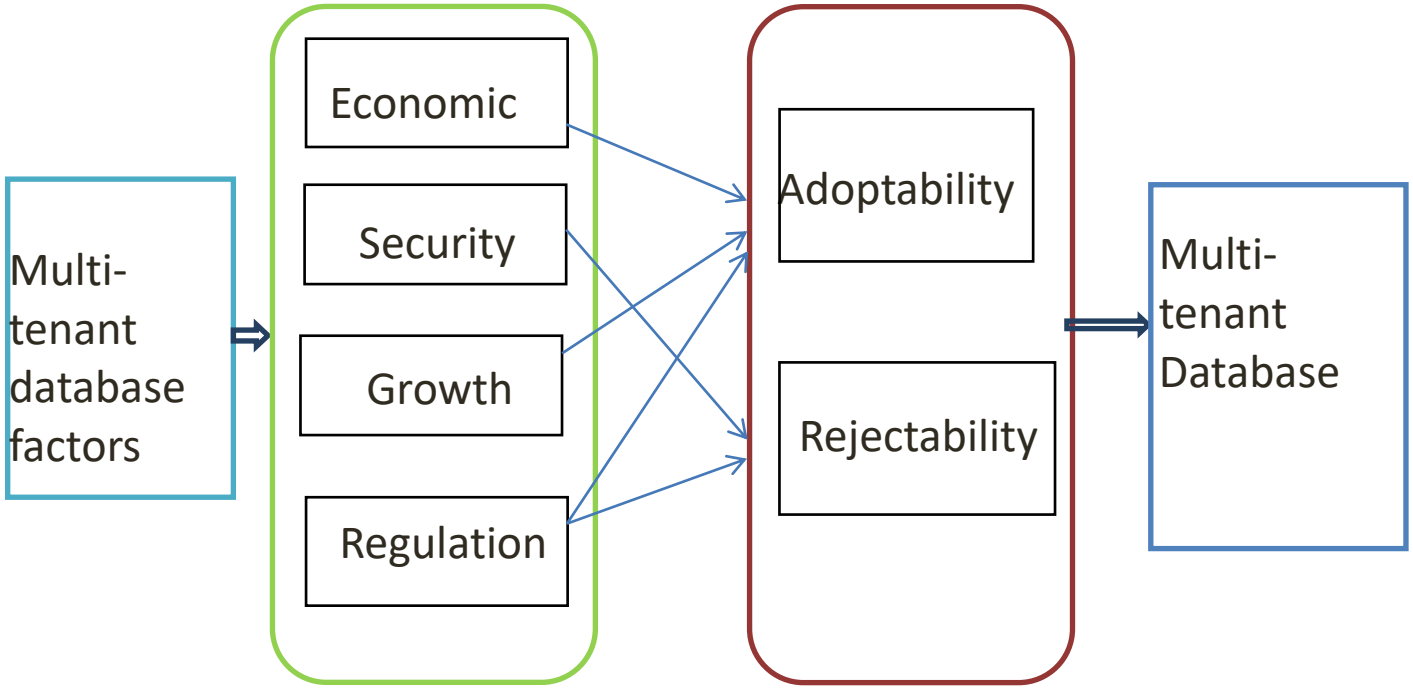


Figure 4.1- Multi-tenant Conceptual Framework

4.5 Summary

This Chapter has shown that multi-tenant database helps reduce the cumulative cost that will be incurred by organisations on the development and deployment of an on premises database system. It has also shown that economic factors can constrain the development attempt which obviously can influence the choice of appropriate approach of multi-tenant database models. One need to take into consideration the growth rate of your database which include the number of tenants, the number of users per tenant, size of each tenant database, growth rate of tenants and lastly the growth rate of tenant database. The different security issues related to multi-tenant database which include data isolation, scalability, flexibility and customization has been presented in the Chapter. Effort must be made to look into the regulatory provisions for prospective tenants. The regulatory laws of all the tenants on a particular multi-tenanted database must be the same in order to have harmony in the laws that govern the system. Finally, a framework was developed based on the hypotheses showing the direction each of the factors tends towards based on the literature reviewed. The next Chapter will show the type of data collected, method of data collection and type of analysis carried out.

CHAPTER FIVE

5.0 Statistical Analysis

5.1 INTRODUCTION

The Chapter Four of this thesis discussed the research methodology undertaken for this research including research approach, research design, research strategies, data collection methods and data analysis methods.

This questionnaire analysis Chapter is the data analysis which deals with the quantitative data. Results of different sections of the questionnaire are shown in tables and diagrams and explanations are given for each. Two stages of administration of the questionnaire were carried out and the discussions are provided based on the stages. These two stages are listed below:

i. Analysis on focus group.
ii. Analysis on the survey.

5.2 Data Analysis on the Focus Group

Data were collected from the Oracle User Group which is the first stage of the questionnaire administration.

5.2.1 The Focus Group.

The UK Oracle user group (UKOUG) members were the participants in this survey. This is an association formed since 1983 with the mission to serve the Oracle community. This is based in United Kingdom and made up of over 8,500 people working for a variety of Oracle customers, Oracle partners and Oracle corporations. They come together every year with the purpose of education, innovation and information. The survey was conducted in the December, 2013 conference of the organisation held in Manchester. For ethical reasons, all respondents' personal details were made anonymous. This group of people were selected because of their high level of involvement with databases and also because of the recent product of Oracle called 12C which is based on the multitenant concept. More than 100 questionnaires were administered but only 30 were returned and carefully analysed to obtain the outcome of this stage. The reasons for the few responses were either due to participant's

company policy or in some situations participants were reluctant to complete the questionnaire for reasons of data protection.

5.2.2 Questionnaire Structure.

As discussed in the Methodology chapter, the questionnaire has been divided into two main sections

1. Respondent's general information.
2. Investigation of Multi-tenant Database drive.

A copy of the questionnaire has been provided in Appendix A, and if necessary some sections of it are shown in the chapter to avoid any ambiguities while discussing the details. The results extracted from each of these sections listed above are explained below together with their related diagrams and tables to illustrate the results clearly.

5.2.3 Analysis

This section covers the analysis on the first survey carried out with the oracle user group in UK. The number of questionnaire administered was 100 with 30 returned which represent a figure of 30% response rate. This low rate was as results of participants' organisational policies. And only one contact was allowed by the organisers of the conference to administer the questionnaire. And according to Romanov and Nir (2010) in Fuch et al (2013) demonstrated that limiting the number of contact attempts resulted in a reduced response rate. There was no other opportunity to meet the participants another time because of policies in place. According to Gunduz et al (2015), all the factors were given equal weight in the questionnaire. The analysis here is based on the question 13 of the questionnaire which is centred on the degree of influence that each of the factors has on the drive towards MTD concept. In analysing question 13 will help target towards the aim of the research. The graphical representation of the responses based on each factor is shown below. The graph shows the lines of the rankings in relation with each factor.

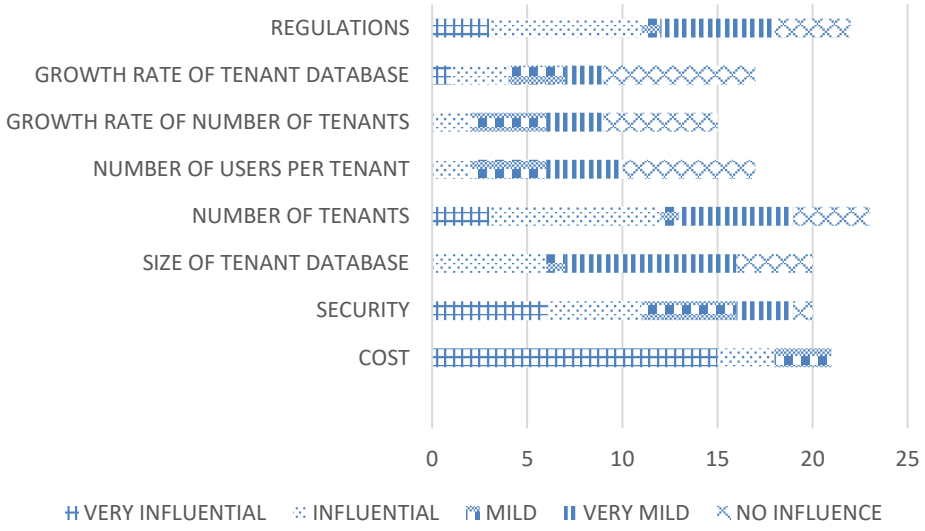


Figure 5.1 – Graphical summary of respondents' data

The graph above shows that cost has the highest value for very influential with 15 respondents agreed for cost as the most influential factor. And none of the respondent indicates that cost has no influence. The growth rate of number of tenants has the highest value for no influence with 8 respondent agreed that it is of no influence to the adoption of MTD. All the factors (1-5) above that makes the growth factor has values for no influence and when put together it shows that growth might has the least effect on the decision to adopt MTD.

Below Table 5.1 illustrates the responses based on each factor, showing the rankings with their frequencies for each of the factors.

Table 5.1 – Summary of Respondents Data

RANKING / FACTORS	VERY INFLUENTIAL	INFLUENTIAL	MILD	VERY MILD	NO INFLUENCE
COST	15	3	3	0	0
SECURITY	6	5	5	3	1
SIZE OF TENANT DATABASE	0	6	1	9	4
NUMBER OF TENANTS	3	9	1	6	4
NUMBER OF USERS PER TENANT	0	2	4	4	7
GROWTH RATE OF NUMBER OF TENANTS	0	2	4	3	6
GROWTH RATE OF TENANT DATABASE	1	3	3	2	8
REGULATIONS	3	8	1	6	4

In analysing this data further, a simple quantitative statistical method was adopted known as weighted score method also known as numerical indicator (Abeysekera 2001 p10).

Very influential – 5, influential – 4, Mild – 3, very mild – 2, No influence 1

Using Weighted Score to analyse the results from the 30 respondents for each factor.

$$\text{The formula} = \frac{N_1 * R_1 + N_2 * R_2 + N_3 * R_3 \ldots\ldots + N_N * R_N}{N_1 + N_2 + N_3 \ldots\ldots + N_N} \qquad (1)$$

Where N = the frequency of each ranking.

R = the ranking value.

The above formula (1) is applied for Cost as shown below

$$\text{Cost} = \frac{(15*5) + (3*4) + (3*3) + (0*2) + (0*1)}{15 + 3 + 3 + 0 + 0} = 4.57$$

After using the above formula (1), the results are represented in the Table 5.2 below. The Table 5.3 below shows the results for the factors after re-grouping them into four. Growth value in Table 5.3 is the average value of all the factors that make up the growth as explained above.

Table 5.2 - Initial Table of Results

Factor	Result
Cost	4.57
Security	3.6
Size of Tenant Database	2.45
Number of Tenants	3.04
Number of Users per Tenant	2.05
Growth rate of Number of Tenants	2.13
Growth rate of Tenant's Database	2.24
Regulations	3.0

Table 5.3-Finaltable of Results

Factor	Result
Cost	4.57
Security	3.6
Growth	2.38
Regulations	3.0

5.2.4. Findings and Discussion

The respondents' data shows that all factors were ranked but not all respondents ranked all the factors. This is because the respondents said they are not sure if those factors really have impact or influence on the decision towards MTD. However, looking at the result of the analysis, cost has the highest indicator value of 4.57 which means it has the greatest impact based on the perspective of the respondents. This means that cost should be first factor to be considered. This is the total cost of ownership which includes several cost as explained above. Since every organisation is set up for business purpose that means that any form of saving cost and increasing revenue will inform the decision to be taking.

Security is another factor that has second indicator value of 3.6 and should be consider after cost has been taking into consideration. Data safety remains one major aspect any organisation would like to maintain. The risk of data being exposed to another party must be guided against in MTD so that perspective tenant will be confident with the service provided. There should be a good service level agreement to protect your data policies.

Regulations have a value of 3.0 which means that this is consider even before the growth factor. Regulations that govern the environment of your tenants as a service provider must be considered. Tenants should be conversant with the laws and regulations that protect your database before even considering the MTD service provider to engage. These regulations differs based on regions, continent, nations or state where tenants are located.

Growth which includes the following factors; the size of tenant database, number of tenants, number of users per tenant, growth rate of number of tenants and growth rate of tenant's database has an average value of 2.38 shows that is the least factor that should be considered. This might suggest that it will have the least degree of impact on the decision about MTD.

5.2.5. Conclusion

This research has shown that MTD will enable service provider to transform a fixed service to a variable services of their technology footprint to enhance business agility, optimize their operations and lower their operational cost thus drive business. Also, this research has been able to identify the factors and proved that these identified factors have real impact on the adoption of MTD. These set of responses from the participants were able to prove the impact of these factors through their contribution in this research. However, this stage of the research has not proven whether it is positive or negative impact that is whether it drives towards the adoption or away from the adoption of the concept.

This research has also shown that MTD helps reduce the TCO involved in acquiring a dedicated database system. The growth of the MTDs include the number of tenants, the number of users per tenant, the size of each tenant's database, growth rate of tenants and the growth rate of tenant's database should all be a consideration in adopting MTD. Technical issues like data isolation, scalability, flexibility and customization are all incorporated in the security consideration. Finally, the regulations that govern all tenants in a MTD must be made harmonised for adherence to the laws.

5.3 Data Analysis on the survey.

This is main section for the data analysis which covers data survey from a set of expert in the field of database. Questionnaires were administered online and across the whole world response were received from different continent of the world which includes; America, Europe, Asia, Africa and Australia. There are a total of 41 participants in the survey. The

reason for low participation has to do majorly with the level of technicality of the questionnaire since the response is expected from a certain set of experts in the concept of MTD. This is also in accordance to Bryman (2001) said that survey is an appropriate means of collecting data when the information sought is reasonably specific and familiar to the respondents. The survey is specific to experts in MTD. Despite the important role of surveys in innovation studies, relatively little attention is given to the challenges of achieving high response rates. Survey participation is a particularly acute issue for web surveys, which tend to suffer from lower response rates than other survey modes. According to (Couper, 2000, Fricker et al., 2005, Kaplowitz et al., 2004 and Rogelberg and Stanton, 2007) in Sauermann and Roach (2013), short and direct surveys involving phone follow-ups can achieve relatively high response rates of 40–70% more detailed online surveys often exhibit lower response rates of around 10–25%. This is one other reason for this low response rate in this research. From the percentage distribution Table 5.4 shows the credibility of respondents as they are all majorly IT expert with database administrators having the highest number. Also, the response shows their professions, type of organisation and email addresses. And the email addresses are all indicating their true affiliation to their claimed organisation.

5.3.1 Questionnaire Structure

Clearly in this part of the survey, time was considered to a very great extent. Hence bearing this in mind, the questionnaire was constructed of multiple choice and closed questions. As used in section 5.2.2 above which is in line with the methodology chapter, the questionnaire has been divided into two main sections

1. Respondent's general information - This comprises of background questions about the respondents and their organisational information.

2. Investigation of Multi-tenant Database drive – This comprises questions about the concept of MTD and their opinion about the drive towards it.

A copy of the questionnaire has been provided in Appendix C, and if necessary some sections of it are shown in the chapter to avoid any ambiguities while discussing the details.

5.3.2 Method of Analysis.

This research on MTD is largely quantitative and is concerned with measurement of majorly the nominal and ordinal variables. The data from the survey were coded into SPSS and represented in numerical values. These data were subjected to the following statistical tools;

1. Percentage Frequency Distribution.
2. Relative Importance Index (RII).
3. Cross Tabulation.

Each of these is explained in detail in the next sections and how they were used for the purpose of this study.

5.4 Percentage Frequency Distribution (PFD).

The percentage frequency distribution is statistical table or diagram that which exhibits a simple analysis of statistical data in terms of percentages, this form of diagram could vary from a bar chart and a pie chart (Dodge 2006 p306). The frequency distribution is always the total frequency equated to 100 and the individual class frequencies expressed in proportion to that figure. The frequency of a particular observation is the number of times the observation occurs in the data. This means that frequency distribution can show either the actual number of observations falling in each range or the percentage of observations. Therefore percentage frequency distribution is the representation of data to show both the number of times the observation occurs and the percentage of the occurrence of the observation.

5.4.1 PFD Application and Results

The results extracted from each of the sections listed above from the questionnaire are explained below together with related tables to illustrate the results clearly using percentage frequency distribution table while the bar charts presented in the appendix F. The questionnaires have been completed by different categories of IT experts. These ranges from the following group of professions;

1. Database Administrators.
2. IT Consultants.
3. Researchers.
4. IT Engineers.

5. Programmers.
6. Database Users.
7. Systems Administrators.
8. IS Managers.
9. Software Engineers.

The demography of each category is explained below regarding their age, type of profession, type of organisation and country. The overall number of participants returning valid online questionnaire is 39 out of the 41.

Table 5.4 Percentage Distribution of Respondent' Profession

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unknown	1	2.6	2.6	2.6
	Database Administrator	16	41.0	41.0	43.6
	IT Consultant	4	10.3	10.3	53.8
	Researcher	5	12.8	12.8	66.7
	IT Engineer	3	7.7	7.7	74.4
	Programmer	1	2.6	2.6	76.9
	Database User	2	5.1	5.1	82.1
	Systems Administrator	2	5.1	5.1	87.2
	Information System Manager	2	5.1	5.1	92.3
	Software Engineer	3	7.7	7.7	100.0
	Total	39	100.0	100.0	

Table 5.4 shows that Database Administrator was the dominant profession (41%) who participated in the survey, this was followed by the researcher profession (12.8%) and the least was those from the programmer and the unknown profession (2.6%).

Table 5.5 Percentage distribution of respondents' age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under 30	7	17.9	17.9	17.9
	30 -39	18	46.2	46.2	64.1
	40 – 49	13	33.3	33.3	97.4
	50 and Above	1	2.6	2.6	100.0
	Total	39	100.0	100.0	

Table 5.5 shows that 30-40 age group has the highest percentage value of the respondents with 46.2%, this was followed by 40-49 age group with 33.3%, there after the under 30 group with 17.9% and the least was the 50 and above with 2.6%.

Table 5.6 Percentage Distribution of respondents' Country

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Australia	2	5.1	5.1	5.1
	UK	10	25.6	25.6	30.8
	USA	4	10.3	10.3	41.0
	Canada	4	10.3	10.3	51.3
	France	1	2.6	2.6	53.8
	Germany	2	5.1	5.1	59.0
	Ghana	1	2.6	2.6	61.5
	India	1	2.6	2.6	64.1
	Nigeria	9	23.1	23.1	87.2
	South Africa	3	7.7	7.7	94.9
	South Korea	1	2.6	2.6	97.4
	Unknown	1	2.6	2.6	100.0
	Total	39	100.0	100.0	

From Table 5.6 the highest percentage value for the respondent country is the UK with 25.6%, this was closely followed by Nigeria with 23.1% and the least was those with 2.6% which include France, Ghana, India, South Korea and the unknown country.

Table 5.7 Percentage Distribution of Organisation Types

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unknown	1	2.6	2.6	2.6
	IT Service Provider	13	33.3	33.3	35.9
	Software Developer	1	2.6	2.6	38.5
	IT Consultancy	3	7.7	7.7	46.2
	IT Service Provider and Consultancy	3	7.7	7.7	53.8
	Educational	8	20.5	20.5	74.4
	UK Local Authority	1	2.6	2.6	76.9
	Oil and Gas	4	10.3	10.3	87.2
	Internet Service Provider	1	2.6	2.6	89.7
	Database Service Provider	3	7.7	7.7	97.4
	Estate Management	1	2.6	2.6	100.0
	Total	39	100.0	100.0	

Table 5.7 shows that the type of organisation of respondents with the highest percentage participation value is the IT Service Providers with 33.3%, this was followed by the Educational sector with 20.5%. And the least was those with 2.6% which include the unknown, Software Developer, UK Local Authority, Internet Service Provider and Estate Management.

Table 5.8 The Percentage Distribution of the Use of database system

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	38	97.4	97.4	97.4
	No	1	2.6	2.6	100.0
	Total	39	100.0	100.0	

Table 5.8 shows that percentage and numerical distribution of the use of database system by the respondents. This shows that 97.4% of the respondents use database system while only 2.6% does not make use of database.

Table 5.9 The Percentage Distribution of Level of involvement with database

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Novice(less than 6 Months)	1	2.6	2.6	2.6
	Intermediate(between 6month -3 years)	9	23.1	23.7	26.3
	Expert (3 years above)	28	71.8	73.7	100.0
	Total	38	97.4	100.0	
Missing	System	1	2.6		
Total		39	100.0		

Table 5.9 shows the percentage distribution of the level of involvement with database system. This shows that 2.6% of the respondents were Novice, 23.7% intermediate those who have just used database system for 6 months to 3 years and 73.7% for the expert who have used the database system for more than 3 years.

Table 5.10 Percentage Distribution of the Database Administrators

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	22	56.4	59.5	59.5
	No	15	38.5	40.5	100.0
	Total	37	94.9	100.0	
Missing	System	2	5.1		
Total		39	100.0		

Table 5.10 shows the percentage distribution of database administrators among the respondents. This shows that 37 out the 39 respondents participated in this question where 59.5% represents those that are database administrator and 40.5% for those that are not administrator.

Table 5.11 Percentage Distribution of Respondent's Awareness of MTD

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	37	94.9	94.9	94.9
	No	2	5.1	5.1	100.0
	Total	39	100.0	100.0	

Table 5.11 shows the distribution of respondent's awareness of MTD concept. There is 100% participation here where 94.9% for those that aware of MTD while only 5.1% for those not aware of the concept.

Table 5.12 Percentage Distribution of MTD usage

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	26	66.7	70.3	70.3
	No	11	28.2	29.7	100.0
	Total	37	94.9	100.0	
Missing	System	2	5.1		
Total		39	100.0		

Table 5.12 shows the percentage distribution of the usage of MTD where 37 out of 39 respondents participated in this question. There is a value of 66.7% for those that have used MTD and 28.2% for those that have never used MTD before.

Table 5.13 Percentage Distribution of number of Tenants on MTD used

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Below 10	8	20.5	30.8	30.8
	11-20	9	23.1	34.6	65.4
	21-30	6	15.4	23.1	88.5
	31-40	3	7.7	11.5	100.0
	Total	26	66.7	100.0	
Missing	System	13	33.3		
Total		39	100.0		

Table 5.13 shows the percentage distribution for the range of number of tenant in the MTD used by these respondents. This shows that 26 out of 39 participated in this question. Below 10 tenants is 30.8%, 11-20 has the value of 34.6%, 21-30 has the value of 23.1% while 31-40 has the value of 11.5%. This shows that 11-20 range has the highest value while 31-40 has the least value.

Table 5.14 Percentage Distribution MTD provider or user

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Provider	11	28.2	42.3	42.3

	User	15	38.5	57.7	100.0
	Total	26	66.7	100.0	
Missing	System	13	33.3		
Total		39	100.0		

Table 5.14 shows the percentage distribution between those that are MTD providers and users where 26 out of 39 respondents participated. Providers have the value of 42.3% while Users are 57.7%. This indicates that users are more than the provider among the respondents.

Table 5.15 Percentage Distribution of Time as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	28	71.8	75.7	75.7
	No	9	23.1	24.3	100.0
	Total	37	94.9	100.0	
Missing	System	2	5.1		
Total		39	100.0		

Table 5.15 shows the percentage distribution of respondent's view about time as a factor where 37 out of 39 participated in this question. Yes has a value of 75.7%, and No has 24.3%.

Table 5.16 Percentage Distribution of Cost as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	35	89.7	97.2	97.2
	No	1	2.6	2.8	100.0
	Total	36	92.3	100.0	
Missing	System	3	7.7		
Total		39	100.0		

Table 5.16 shows the percentage distribution of respondent's view about cost as a factor with 36 out of 39 participated in this question. This shows that Yes has a value of 97.2%, No has 2.8%.

Table 5.17 Percentage Distribution of Data Isolation as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	29	74.4	80.6	80.6
	No	7	17.9	19.4	100.0

	Total	36	92.3	100.0	
Missing	System	3	7.7		
Total		39	100.0		

Table 5.17 shows the percentage distribution of respondent's view about data isolation as a factor with 36 out 39 participated in this question. This shows that Yes has a value of 80.6%, No has 19.4%.

Table 5.18 Percentage Distribution of Scalability as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	33	84.6	94.3	94.3
	No	2	5.1	5.7	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.18 shows the percentage distribution of respondent's view about Scalability as a factor with 35 out of 39 participated. This shows that Yes has a value of 94.3%, No has 5.7%.

Table 5.19 Percentage Distribution of Flexibility as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	27	69.2	77.1	77.1
	No	8	20.5	22.9	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.19 shows the percentage distribution of respondent's view about Flexibility as a factor with 35 out of 39 participated. This shows that Yes has a value of 77.1%, No has 22.9%.

Table 5.20 Percentage Distribution of Customization as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	29	74.4	82.9	82.9
	No	6	15.4	17.1	100.0

	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.20 shows the percentage distribution of respondent's view about Flexibility as a factor with 35 out of 39 respondents participated. This shows that Yes has a value of 82.9%, No has 17.1%.

Table 5.21 Percentage Distribution of Regulation as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	26	66.7	74.3	74.3
	No	9	23.1	25.7	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.21 shows the percentage distribution of respondent's view about Regulation as a factor with 35 out of 39 respondents participated. This shows that Yes has a value of 74.3% and No has 25.7%.

Table 5.22 Percentage Distribution of Size of Tenant DB as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	25	64.1	71.4	71.4
	No	10	25.6	28.6	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.22 shows the percentage distribution of respondent's view about Size of Tenant's Database as a factor with 35 out of 39 respondents participated. This shows that Yes has a value of 71.4% and No has 28.6%.

Table 5.23 Percentage Distribution of Number of Tenant as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	28	71.8	80.0	80.0
	No	7	17.9	20.0	100.0

	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.23 shows the percentage distribution of respondent's view about Number of Tenants on the MTD as a factor with 35 out of 39 respondent participated. This shows that Yes has a value of 80% and No has 20%.

Table 5.24 Percentage Distribution of Number of User/Tenant as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	28	71.8	80.0	80.0
	No	7	17.9	20.0	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.24 shows the percentage distribution of respondent's view about Number of Users per Tenants as a factor with 35 out 39 participated in this question. This shows that yes has a value of 80% and No has 20%.

Table 5.25 Percentage Distribution of the Growth rate of Number of Tenant as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	20	51.3	57.1	57.1
	No	15	38.5	42.9	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.25 shows the percentage distribution of respondent's view about the Growth rate of number of Tenants as a factor with 35 out of 39 respondents participated. This shows that Yes has a value of 57.1% and No has 42.9%.

Table 5.26 Percentage Distribution of Growth rate of Tenant DB as a Factor

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	25	64.1	71.4	71.4

	No	10	25.6	28.6	100.0
	Total	35	89.7	100.0	
Missing	System	4	10.3		
Total		39	100.0		

Table 5.26 shows the percentage distribution of respondent's view about the Growth rate Tenants Database as a factor with 35 out of 39 respondents participated. This shows that Yes has a value of 71.4% and No has 28.6%.

5.4.2 Discussion on Findings from PFD

The PFD analysis has helped show the respondents' profession, age range, country, level of database usage, MTD awareness level, MTD usage level etc. For the purpose of this study, concentration is on the data relevant to the aim and objectives of the research.

In Table 5.9, the total number of participant is 38 with 73.7% of this are the experts that have used database for more than 3 years. While for the intermediate users which are those who have used database for a period ranging between 6 months to 3 years has a value of 23.7%. Table 5.10 also shows that 59.5% of the total respondents were the database administrators while 40.5% are not but most of which are also IT related job descriptions. This is also proven by the results in table 5.4 with 2.6% non-IT profession while the rest 97.4% are all IT related profession out of which 41% are majorly database administrators. These three PFD tables illustrate that a very high percentage of the respondents are expert in the field and they are very much involved in the database management.

The results in Table 5.11 are further expatiated by Table 5.12 where 94.9% of the participants are aware of the MTD concept. This is a strong indication that participants are much conversant with the subject of study. Table 5.12 shows the number and percentage of the respondents who not only are aware but also make use of MTD is very high with a value of 66.7%. This means that 66.7% makes use of MTD in their operations. Therefore, the results and conclusions from this study will be of a very high quality with a sound judgement because of the areas of specialization of the participants. This is a proof of the quality of the research.

While Table 5.13 shows that number of tenants on those with MTD platform with 11-20 tenants carries the highest value of 23.1% with 31-40 tenants has the least value of 7.7%. Among those that make use of MTD, table 6.14 shows that 28% are service provider while

38.5% users. The results in table 5.15 – 5.26 illustrate the percentage distribution of Yes and No responses to all the factors.

5.5 Relative Importance Index (RII)

The use of the term relative importance refers to the contribution a variable makes to the prediction of a criterion variable by itself and in combination with other predictor variables (Johnson and LeBreton 2004 in Tronidandel and LeBreton 2011). This definition considers only the relative contribution of a variable to total predictable variance and makes no assumptions about either the statistical significance or practical significance associated with a particular predictor. RII aims of generating an index that can ordinally arrange those variables being studied in terms of (respondents') perceived agreement, relevance, importance, risk, or other discriminating criterion.

Accordingly to Holt (2013 p5) among the literature the RII may be described differently to reflect its application to a particular study – such as "relative agreement index", "relative importance index", "risk rating index" and so on. Therefore, for the purpose of this research which is based on measuring the impact of some factor on the adoption of MTD, the concept of RII is adopted and can be referred to as RELATIVE IMPACT INDEX.

There are about eight different RII methods identified and discussed by Holt (2013 p6), but for the purpose of this study the method adopted is explained below which is the method also used by Sambasivan and Soon (2007); Gunduz et al (2013).

$$RII = \frac{\Sigma W}{A*N} \qquad (2)$$

Where W is the weighting given to each factor by the respondents (ranging from 1 to 5),
A is the highest weight (i.e. 5 in this case), and N is the total number of respondents.

The RII value had a range from 0 to 1 (0 not inclusive), higher the value of RII, more important was the cause or effect (i.e. the more the impact in this case).

5.5.1 RII Application and Results

This section shows the results from the application RII formula above on the data collected to show the degree of impact of each factor that influences the move towards the concept of MTD.

Formula (2) above is applied for respondent's data for Time as a factor

$RII_{(TIME)} = [(3*4) + (4*20) + (5*4)]/5*28$

$= 12 + 80 + 20/140 \quad = 0.80$

The results for others are shown in the Table 5.27 below.

Table 5.27 – Relative Impact Index (RII) Results

ECONOMIC	Weight/Scale					
Factors	1	2	3	4	5	RII
Time	0	0	4	20	4	0.80
Cost	0	0	4	12	19	0.89
GROWTH	Weight/Scale					
Factors	1	2	3	4	5	RII
Size of tenant DB	0	5	14	5	1	0.62
No. of tenants	0	4	15	9	0	0.64
No. of users per tenant	1	11	11	5	0	0.54
Growth rate of No. of tenants	3	7	7	2	1	0.51
Growth rate of tenant DB	3	8	8	5	0	0.53
SECURITY	Weight/Scale					
Factors	1	2	3	4	5	RII
Data Isolation	0	0	7	17	5	0.79
Scalability	0	5	15	12	1	0.65
Flexibility	0	9	9	8	1	0.61
Customization	0	2	15	11	2	0.69
MAIN FACTORS	Weight/Scale					
FACTORS	1	2	3	4	5	RII
ECONOMIC	0	2	0	8	25	0.92
GROWTH	1	1	25	7	0	0.62
SECURITY	1	0	0	26	7	0.82
REGULATIONS	2	12	6	4	2	0.54

5.5.2 Discussion on Findings from RII

This section of discussion is based on the results from the application of RII formulae used in the analysis section. The discussion is based on the re-grouped four factors as itemized in the following sub-headings. This helps to determine the degree of impact of these factors in relation to the MTD concept.

5.5.2.1 Economic Factors

The use of the RII formulae in section 6.5.1 above as shown some important information about the factors as identified. This show the level of impact each of the factor has in adopting the concept of MTD. The value of RII for the time and cost as factors shows the relative impact. Time has a value of 0.80 and cost with 0.89 respectively. This means that

both factors have a great impact on the decision process towards the concept but with cost having greater impact. These two factors form the economic factor which is considered to have the highest RII of 0.92 from the table 5.30 after the regrouping of the factors into four.

5.5.2.2 Security Factors

There are four sub factors that make up the security which includes data isolation, scalability, flexibility and customization. The use of RII formulae in section 6.5.1 above also shows the value for each of these factors. Data isolation as a factor has the highest value with 0.79 followed by customization with 0.69; the next factor in value is the scalability with 0.65 while flexibility is the last in the list with the value of 0.61. These four factors form the security factor which is considered to have second highest RII after economic factor with the value of 0.82.

5.5.2.3 Growth Factors

There are five sub factors here based on the study that make up the growth factor which are the size of tenant's database; number of tenants; the number of users per tenant; growth rate of number of tenants and growth rate of tenant's database. The use of the RII formulae shows that number of tenants has the highest RII value of 0.64 followed by the size of tenant database with the value of 0.62. While number of users per tenant, growth rate of tenant database and growth rate of number of tenants has the values of 0.54, 0.53 and 0.51 respectively. These five factors form the growth factor which has the third highest RII value 0.62 from table 5.30.

5.5.2.4 Regulation Factors

Regulation as a factor does not have any other sub factor based on the study. The application of RII formulae shows that it does have the least RII value of 0.54 from table 5.27.

In conclusion, this stage of the analysis method has shown that the economic has the highest impact followed by security, after security we have growth and after growth is regulation.

5.6 Cross-Tabulation.

Cross Tabulation also known as Contingency table is a way of summarising the association between variables that have nominal or categorical data (Hinton et al 2014). We use frequency tables to summarize a single categorical variable while we use a cross-tabulation to summarize the relationship between two categorical variables. A cross-tabulation (or *crosstab* for short) is a table that depicts the number of times each of the possible category combinations occurred in the sample data (Miller and Acton 2009). The crosstabs procedure can use numeric or string variables defined as nominal, ordinal, or scale. However, crosstabs should only be used when there are a limited number of categories.

A cross tabulation table shows the combinations of results of different questions of a survey in a table with the results of one question as the rows and the results of another question as the columns. The frequency data can be represented in a table with the rows as the conditions of one variable and the columns as the conditions of a second variable (Hinton et al 2014). This method is therefore adopted in this research to examining the relationship of responses of participants between different variable within the survey. The results of the cross tabulations are presented in the next section of this thesis.

5.6.1 Cross-Tabulation Results and Discussions

This section shows the results of applying cross tabulation to some of the variables to illustrate the relationships between them. This helps to identify the level participation of respondents to a particular variable. These variables include the level of involvement with database, database administrator, awareness of MTD and MTD usage. Theses variables are cross tabulated against the four major factors. These are shown in the following tables.

Table 5.28 Crosstab of level of involvement * Economic factor towards Adoption/Rejection

			Economic Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Level of involvement with database	Intermediate(between 6month -3 years)	Count	6	2	8
		% within Level of involvement with database	75.0%	25.0%	100.0%
		% within Economic Factor towards Adoption or Rejection	20.0%	40.0%	22.9%
		% of Total	17.1%	5.7%	22.9%
	Expert (3 years above)	Count	24	3	27
		% within Level of involvement with database	88.9%	11.1%	100.0%
		% within Economic Factor towards Adoption or Rejection	80.0%	60.0%	77.1%
		% of Total	68.6%	8.6%	77.1%
Total		Count	30	5	35
		% within Level of involvement with database	85.7%	14.3%	100.0%
		% within Economic Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	85.7%	14.3%	100.0%

The Table 5.28 shows the cross-tab of the level of involvement with database against economic factor. There are two levels of involvement here; the intermediate (those that have been using database between 6 months and 3 years) and the expert (those that have used database for 3 years and above). We have 8 responses under intermediate of which 6 (17.1%) agreed that economic factor leads to adoption while 2 (5.7%) said it leads to rejection. While under the experts, there are 27 respondents, 24(68.6%) agreed that it leads to adoption while 3(8.6%) said it lead to rejection. Putting the two categories together we have 30 (85.7%) responses for adoption and 5 (14.3%) for rejection. This show that there a very high gap

between the adoption and the rejection. This means that the economic factor has a positive impact and it leads to adoption of MTD. This is just based on the level of involvement with database.

Table 5.29 Crosstab of level of involvement * Regulation Factor towards Adoption/Rejection

			Regulation Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Level of involvement with database	Intermediate(between 6month -3 years)	Count	6	2	8
		% within Level of involvement with database	75.0%	25.0%	100.0%
		% within Regulation Factor towards Adoption or Rejection	23.1%	25.0%	23.5%
		% of Total	17.6%	5.9%	23.5%
	Expert (3 years above)	Count	20	6	26
		% within Level of involvement with database	76.9%	23.1%	100.0%
		% within Regulation Factor towards Adoption or Rejection	76.9%	75.0%	76.5%
		% of Total	58.8%	17.6%	76.5%
Total		Count	26	8	34
		% within Level of involvement with database	76.5%	23.5%	100.0%
		% within Regulation Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%

The Table 5.29 shows the cross-tab of the level of involvement with database against regulation factor towards adoption or rejection. Also, we have 8 responses under the intermediate of which 6 (17.6%) agreed that regulation factor leads to adoption while 2 (5.9%) said it leads to rejection. The expert category, there are 26 respondents, 20 (58.8%) agreed that it lead to adoption while 6 (17.6%) said it leads to rejection. The combination of the two categories amounts to 76.5% for adoption and 23.5% for rejection. This also has

shown that there is a wide gap between adoption and rejection. This means that regulation factor has a positive impact and it leads to the adoption of the concept.

Table 5.30 Crosstab of level of involvement * Security Factor towards Adoption/Rejection

			Security Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Level of involvement with database	Intermediate(between 6month -3 years)	Count	6	2	8
		% within Level of involvement with database	75.0%	25.0%	100.0%
		% within Security Factor towards Adoption or Rejection	40.0%	10.5%	23.5%
		% of Total	17.6%	5.9%	23.5%
	Expert (3 years above)	Count	9	17	26
		% within Level of involvement with database	34.6%	65.4%	100.0%
		% within Security Factor towards Adoption or Rejection	60.0%	89.5%	76.5%
		% of Total	26.5%	50.0%	76.5%
Total		Count	15	19	34
		% within Level of involvement with database	44.1%	55.9%	100.0%
		% within Security Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	44.1%	55.9%	100.0%

The cross-tabulation of level of involvement with database against the security factor shows some results from Table 5.30. There are 8 intermediate respondents of which 6 (17.6%) agreed that security leads to adoption while 2 (5.9%) said it leads to rejection. Under the expert category there is a deviating result from the intermediate. There are 26 respondents, 9 ((26.5%) agreed that security leads to adoption while 17 (50%) said it leads to rejection. The combination of these two categories gives a total result of 15 (44.1%) for adoption while 19 (55.9%) for rejection. In this result, there is no much gap between adoption and rejection but

since more of the experts agreed that security leads to rejection, we can conclude that security will not encourage adoption of MTD. This means that security is more of negative impact than positive.

Table 5.31 Crosstab of level of involvement * Growth Factor towards Adoption/Rejection

			Growth Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Level of involvement with database	Intermediate(between 6month -3 years)	Count	6	2	8
		% within Level of involvement with database	75.0%	25.0%	100.0%
		% within Growth Factor towards Adoption or Rejection	23.1%	25.0%	23.5%
		% of Total	17.6%	5.9%	23.5%
	Expert (3 years above)	Count	20	6	26
		% within Level of involvement with database	76.9%	23.1%	100.0%
		% within Growth Factor towards Adoption or Rejection	76.9%	75.0%	76.5%
		% of Total	58.8%	17.6%	76.5%
Total		Count	26	8	34
		% within Level of involvement with database	76.5%	23.5%	100.0%
		% within Growth Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%

The Table 5.31 is the cross-tabulation result of level of involvement with database against growth factor towards adoption or rejection. There are 8 intermediate respondents of which 6 (17.6%) agreed that growth as a factor leads to adoption while 2 (5.9%) said it lead to rejection. The results shows that there 26 expert respondents of which 20 (58.8%) agreed that growth leads to adoption while 6 (17.6%) said it lead to rejection. The combination of the two categories gives a total result of 26 (76.5%) for adoption while 8 (23.5%) for rejection. This

shows that there is a wide gap between the adoption and the rejection. Therefore conclusion can be made that growth factor encourages that adoption of MTD. This also means that growth has a positive impact.

Table 5.32 Crosstab of Are you a Database Administrator * Economic Factor towards adoption/Rejection

			Economic Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Are you a Database Administrator	Yes	Count	18	4	22
		% within Are you a Database Administrator	81.8%	18.2%	100.0%
		% within Economic Factor towards Adoption or Rejection	60.0%	80.0%	62.9%
		% of Total	51.4%	11.4%	62.9%
	No	Count	12	1	13
		% within Are you a Database Administrator	92.3%	7.7%	100.0%
		% within Economic Factor towards Adoption or Rejection	40.0%	20.0%	37.1%
		% of Total	34.3%	2.9%	37.1%
Total		Count	30	5	35
		% within Are you a Database Administrator	85.7%	14.3%	100.0%
		% within Economic Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	85.7%	14.3%	100.0%

There is a cross-tabulation of database administrator against the economic factor shown in Table 5.32. The result in this table shows that there are 22 respondents that said yes they are database administrator out of which 18 (51.4%) of them agreed that economic factor leads to adoption while 4 (11.4%) said it lead to rejection. While those that said no (not database administrator) are 13 respondents out of which 12 (34.3%) said it leads to adoption while only 1(2.9%) said it leads to rejection. The combination of these two categories of participants gives a total result of 30 (85.7%) for adoption while 5 (14.3%) for rejection. This is a wide margin between the adoption and the rejection. Conclusion can therefore be made that economic factor will lead to adoption of MTD and that economic factor has a positive impact on the drive towards the concept.

Table 5.33 Crosstab of Are you a Database Administrator * Regulation Factor towards adoption/Rejection

			Regulation Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Are you a Database Administrator	Yes	Count	16	5	21
		% within Are you a Database Administrator	76.2%	23.8%	100.0%
		% within Regulation Factor towards Adoption or Rejection	61.5%	62.5%	61.8%
		% of Total	47.1%	14.7%	61.8%
	No	Count	10	3	13
		% within Are you a Database Administrator	76.9%	23.1%	100.0%
		% within Regulation Factor towards Adoption or Rejection	38.5%	37.5%	38.2%
		% of Total	29.4%	8.8%	38.2%
Total		Count	26	8	34
		% within Are you a Database Administrator	76.5%	23.5%	100.0%
		% within Regulation Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%

The Table 5.33 shows the cross-tabulation of database administrator against regulation factor towards adoption or rejection of MTD. There are 21 respondents that said yes they are administrator out of which 16 (47.1%) of them agreed that regulation leads to adoption while 5 (14.7%) said it leads to rejection. The result in the table also shows that 13 respondents said no (not database administrator) out of which 10 (29.4%) said it leads to adoption while 3 (8.8%) said it leads to rejection. The combination of these two categories of participants gives a total result of 26 (76.5%) for adoption while 8 (23.5%) for rejection. This also has shown a wide margin between adoption and rejection. This is indicative of positive impact toward the drive about MTD. Therefore, regulation can lead to the adoption of the concept.

Table 5.34 Crosstab of Are you a Database Administrator * Security Factor towards Adoption/Rejection

			Security Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Are you a Database Administrator	Yes	Count	9	12	21
		% within Are you a Database Administrator	42.9%	57.1%	100.0%
		% within Security Factor towards Adoption or Rejection	60.0%	63.2%	61.8%
		% of Total	26.5%	35.3%	61.8%
	No	Count	6	7	13
		% within Are you a Database Administrator	46.2%	53.8%	100.0%
		% within Security Factor towards Adoption or Rejection	40.0%	36.8%	38.2%
		% of Total	17.6%	20.6%	38.2%
Total		Count	15	19	34
		% within Are you a Database Administrator	44.1%	55.9%	100.0%
		% within Security Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	44.1%	55.9%	100.0%

There is a cross-tabulation of database administrator against the security factor shown in table 5.34. The result in this table shows that there are 21 respondents that said yes they are database administrator out of which 9 (26.5%) of them agreed that security factor leads to adoption while 12 (35.3%) said it lead to rejection. While those that said no (not database administrator) are 13 respondents out of 6 (17.6%) said it leads to adoption while 7 (20.6%) said it leads to rejection. The combination of these two categories of participants gives a total result of 15 (44.1%) for adoption while 19 (55.9%) for rejection. In this result, there is no much wide margin between the adoption and the rejection but more of the respondent said security will lead to rejection. Therefore, security will have a negative impact on the adoption of MTD.

Table 5.35 Crosstab of Are you a Database Administrator * Growth Factor towards Adoption/Rejection

			Growth Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Are you a Database Administrator	Yes	Count	19	2	21
		% within Are you a Database Administrator	90.5%	9.5%	100.0%
		% within Growth Factor towards Adoption or Rejection	73.1%	25.0%	61.8%
		% of Total	55.9%	5.9%	61.8%
	No	Count	7	6	13
		% within Are you a Database Administrator	53.8%	46.2%	100.0%
		% within Growth Factor towards Adoption or Rejection	26.9%	75.0%	38.2%
		% of Total	20.6%	17.6%	38.2%
Total		Count	26	8	34
		% within Are you a Database Administrator	76.5%	23.5%	100.0%
		% within Growth Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%

The Table 5.35 shows the cross-tabulation of database administrator against growth factor towards adoption or rejection of MTD. There are 21 respondents that said yes they are administrator out of which 19 (55.9%) of them agreed that growth leads to adoption while 2 (5.9%) said it leads to rejection. The result in the Table also shows that 13 respondents said no (not database administrator) out of which 7 (20.6%) said it leads to adoption while 6 (17.6%) said it leads to rejection. The combination of these two categories of participants gives a total result of 26 (76.5%) for adoption while 8 (23.5%) for rejection. This also has shown a wide margin between adoption and rejection. This is indicative of positive impact toward the drive about MTD. Therefore, growth can lead to the adoption of the concept.

Table 5.36 Crosstab of Awareness of MTD * Economic Factor towards Adoption/Rejection

			Economic Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Awareness of MTD	Yes	Count	30	5	35
		% within Awareness of MTD	85.7%	14.3%	100.0%
		% within Economic Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	85.7%	14.3%	100.0%
Total		Count	30	5	35
		% within Awareness of MTD	85.7%	14.3%	100.0%
		% within Economic Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	85.7%	14.3%	100.0%

There is a cross-tabulation of awareness of MTD against economic factor shown in Table 5.36. The result shows that there are total numbers of 35 participants who are aware of the concept of MTD out of which 30 (85.7%) of them agreed that economic factor leads to adoption of MTD while only 5 (14.3%) said it lead to rejection. This result is a confirmation of the result obtained in the cross-tab of database administrator against economic factor. Conclusion can also be made that economic factor has positive impact on the drive towards MTD and that it will lead to the adoption of the concept.

Table 5.37 Crosstab of Awareness of MTD * Regulation Factor towards Adoption/Rejection

			Regulation Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Awareness of MTD	Yes	Count	26	8	34
		% within Awareness of MTD	76.5%	23.5%	100.0%
		% within Regulation Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%
Total		Count	26	8	34
		% within Awareness of MTD	76.5%	23.5%	100.0%
		% within Regulation Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%

The Table 5.37 shows the cross-tabulation of awareness of MTD against regulation factor towards the adoption or rejection of MTD. The result shows that there are total numbers of 34 participants who are aware of the concept of MTD out of which 26 (76.5%) of them agreed that regulation factor leads to adoption of MTD while only 8 (23.5%) said it lead to rejection. This result is a confirmation of the result obtained in the cross-tab of database administrator against regulation factor. Conclusion can also be made that regulation factor has positive impact on the drive towards MTD and that it will lead to the adoption of the concept.

Table 5.38 Crosstab of Awareness of MTD * Security Factor towards Adoption/Rejection

			Security Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Awareness of MTD	Yes	Count	15	19	34
		% within Awareness of MTD	44.1%	55.9%	100.0%
		% within Security Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	44.1%	55.9%	100.0%
Total		Count	15	19	34
		% within Awareness of MTD	44.1%	55.9%	100.0%
		% within Security Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	44.1%	55.9%	100.0%

There is a cross-tabulation of awareness of MTD against economic factor shown in Table 5.38. The result shows that there are total numbers of 34 participants who are aware of the concept of MTD out of which 15 (44.1%) of them agreed that security factor leads to adoption of MTD while 19 (55.9%) said it leads to rejection. This result is a confirmation of the result obtained in the cross-tab of database administrator against security factor. Conclusion can also be made that security factor has negative impact on the drive towards MTD and that it will lead to the rejection of the concept.

Table 5.39 Crosstab of Awareness of MTD * Growth Factor towards Adoption/Rejection

			Growth Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Awareness of MTD	Yes	Count	26	8	34
		% within Awareness of MTD	76.5%	23.5%	100.0%
		% within Growth Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%
Total		Count	26	8	34
		% within Awareness of MTD	76.5%	23.5%	100.0%

	% within Growth Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
	% of Total	76.5%	23.5%	100.0%

The Table 5.39 shows the cross-tabulation of awareness of MTD against growth factor towards the adoption or rejection of MTD. The result shows that there are total numbers of 34 participants who are aware of the concept of MTD out of which 26 (76.5%) of them agreed that growth factor leads to adoption of MTD while only 8 (23.5%) said it lead to rejection. This result is a confirmation of the result obtained the cross-tab of database administrator against growth factor. Conclusion can also be made that growth factor has positive impact on the drive towards MTD and that it will lead to the adoption of the concept.

Table 5.40 Crosstab of MTD usage * Economic Factor towards Adoption/Rejection

			Economic Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Have you ever use MTD	Yes	Count	21	3	24
		% within Have you ever use MTD	87.5%	12.5%	100.0%
		% within Economic Factor towards Adoption or Rejection	70.0%	60.0%	68.6%
		% of Total	60.0%	8.6%	68.6%
	No	Count	9	2	11
		% within Have you ever use MTD	81.8%	18.2%	100.0%
		% within Economic Factor towards Adoption or Rejection	30.0%	40.0%	31.4%

		% of Total	25.7%	5.7%	31.4%
Total		Count	30	5	35
		% within Have you ever use MTD	85.7%	14.3%	100.0%
		% within Economic Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	85.7%	14.3%	100.0%

There is a cross-tabulation of MTD usage against the economic factor shown in Table 5.40. The result shows that there are 24 respondents that said yes they use MTD out of which 21 (60%) of them agreed that economic factor leads to adoption while 3 (8.6%) said it lead to rejection. While those that said no (don't use MTD) are 11 respondents out of 9 (25.7%) said it leads to adoption while only 2 (5.7%) said it leads to rejection. The combination of these two categories of participants gives a total result of 30 (85.7%) for adoption while 5 (14.3%) for rejection. This is a wide margin between the adoption and the rejection. This result also confirms other cross-tabulations against Economic factor. Conclusion can therefore be made that economic factor will lead to adoption of MTD and that economic factor has a positive impact on the drive towards the concept.

Table 5.41 Crosstab of MTD usage * Security Factor towards Adoption/Rejection

			Security Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Have you ever use MTD	Yes	Count	11	13	24
		% within Have you ever use MTD	45.8%	54.2%	100.0%
		% within Security Factor towards Adoption or Rejection	73.3%	68.4%	70.6%
		% of Total	32.4%	38.2%	70.6%
	No	Count	4	6	10
		% within Have you ever use MTD	40.0%	60.0%	100.0%
		% within Security Factor towards Adoption or Rejection	26.7%	31.6%	29.4%
		% of Total	11.8%	17.6%	29.4%
Total		Count	15	19	34
		% within Have you ever use MTD	44.1%	55.9%	100.0%

	% within Security Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
	% of Total	44.1%	55.9%	100.0%

There is a cross-tabulation of MTD usage against the security factor shown in Table 5.41. The result shows that there are 24 respondents that said yes they use MTD out of which 11 (32.4%) of them agreed that security factor leads to adoption while 13 (38.2%) said it lead to rejection. While those that said no (don't use MTD) are 10 respondents out of 4 (11.8%) said it leads to adoption while 6 (17.6%) said it leads to rejection. The combination of these two categories of participants gives a total result of 15 (44.1%) for adoption while 19 (55.9%) for rejection. In this result, there is no much wide margin between the adoption and the rejection but more of the respondent said security will lead to rejection. Therefore, security will have a negative impact on the adoption of MTD. This result also confirms other cross-tabulations against security factor.

Table 5.42 Crosstab of MTD usage * Growth Factor towards Adoption/Rejection

			Growth Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Have you ever use MTD	Yes	Count	21	3	24
		% within Have you ever use MTD	87.5%	12.5%	100.0%
		% within Growth Factor towards Adoption or Rejection	80.8%	37.5%	70.6%
		% of Total	61.8%	8.8%	70.6%
	No	Count	5	5	10
		% within Have you ever use MTD	50.0%	50.0%	100.0%
		% within Growth Factor towards Adoption or Rejection	19.2%	62.5%	29.4%
		% of Total	14.7%	14.7%	29.4%
Total		Count	26	8	34
		% within Have you ever use MTD	76.5%	23.5%	100.0%
		% within Growth Factor towards Adoption or Rejection	100.0%	100.0%	100.0%

	% of Total	76.5%	23.5%	100.0%

The Table 5.42 shows the cross-tabulation of MTD usage against growth factor towards adoption or rejection of MTD. There are 24 respondents that said yes they use MTD out of which 21 (61.8%) of them agreed that growth leads to adoption while 3 (8.8%) said it leads to rejection. The result also shows that 10 respondents said no (don't use MTD) out of which 5 (14.7%) said it leads to adoption and 5 (14.7%) also said it leads to rejection. The combination of these two categories of participants gives a total result of 26 (76.5%) for adoption while 8 (23.5%) for rejection. This also has shown a wide margin between adoption and rejection. This is indicative of positive impact toward the drive about MTD. Therefore, growth as a factor can lead to the adoption of the concept. This result also confirms other cross-tabulations against growth factor.

Table 5.43 Crosstab of MTD usage * Regulation Factor towards Adoption/Rejection

			Regulation Factor towards Adoption or Rejection		Total
			Adoption	Rejection	
Have you ever use MTD	Yes	Count	20	4	24
		% within Have you ever use MTD	83.3%	16.7%	100.0%
		% within Regulation Factor towards Adoption or Rejection	76.9%	50.0%	70.6%
		% of Total	58.8%	11.8%	70.6%
	No	Count	6	4	10
		% within Have you ever use MTD	60.0%	40.0%	100.0%
		% within Regulation Factor towards Adoption or Rejection	23.1%	50.0%	29.4%
		% of Total	17.6%	11.8%	29.4%
Total		Count	26	8	34
		% within Have you ever use MTD	76.5%	23.5%	100.0%
		% within Regulation Factor towards Adoption or Rejection	100.0%	100.0%	100.0%
		% of Total	76.5%	23.5%	100.0%

The Table 5.43 shows the cross-tabulation of MTD usage against regulation factor towards adoption or rejection of MTD. There are 24 respondents that said yes they use MTD out of which 20 (58.8%) of them agreed that regulation leads to adoption while 4 (11.8%) said it leads to rejection. The result in the table also shows that 10 respondents said no (don't use MTD) out of which 6 (17.6%) said it leads to adoption while 4 (11.8%) also said it leads to rejection. The combination of these two categories of participants gives a total result of 26 (76.5%) for adoption while 8 (23.5%) for rejection. This also has shown a wide margin between adoption and rejection. This is indicative of positive impact toward the drive about MTD. Therefore, regulation as a factor can lead to the adoption of the concept. This result also confirms other cross-tabulations against regulation factor.

Table 5.44 Crosstab of MTD awareness * Acceptance level of MTD

			Acceptance level of MTD today			Total
			No Difference	Improved	Highly Improved	
Awareness of MTD	Yes	Count	5	25	4	34
		% within Awareness of MTD	14.7%	73.5%	11.8%	100.0%
		% within Acceptance level of MTD today	100.0%	100.0%	100.0%	100.0%
		% of Total	14.7%	73.5%	11.8%	100.0%
Total		Count	5	25	4	34
		% within Awareness of MTD	14.7%	73.5%	11.8%	100.0%
		% within Acceptance level of MTD today	100.0%	100.0%	100.0%	100.0%
		% of Total	14.7%	73.5%	11.8%	100.0%

Table 5.44 shows the cross-tabulation of acceptance level of MTD against the MTD awareness. There are 34 respondents in total here of which all are aware of MTD and out which 5 (14.7%) said there is no difference, 25 (73.5%) said there is improved acceptance and 4 (11.8%) said there is highly improved acceptance. This shows that there is a wide

margin between improved and no difference. And putting together the value of improved and highly improved gives a total value of 29 (85.3%) which is an indication that MTD acceptance has improved tremendously in the last 2years.

Table 5.45 The crosstab of MTD usage * Acceptance level of MTD today

			Acceptance level of MTD today			Total
			No Difference	Improved	Highly Improved	
Have you ever use MTD	Yes	Count	4	16	4	24
		% within Have you ever use MTD	16.7%	66.7%	16.7%	100.0%
		% within Acceptance level of MTD today	80.0%	64.0%	100.0%	70.6%
		% of Total	11.8%	47.1%	11.8%	70.6%
	No	Count	1	9	0	10
		% within Have you ever use MTD	10.0%	90.0%	0.0%	100.0%
		% within Acceptance level of MTD today	20.0%	36.0%	0.0%	29.4%
		% of Total	2.9%	26.5%	0.0%	29.4%
Total		Count	5	25	4	34
		% within Have you ever use MTD	14.7%	73.5%	11.8%	100.0%
		% within Acceptance level of MTD today	100.0%	100.0%	100.0%	100.0%
		% of Total	14.7%	73.5%	11.8%	100.0%

Table 5.45 shows the cross-tabulation of acceptance level of MTD against the MTD usage. There are 34 respondents in total here of which some have used MTD and some have not used it before but all are aware of MTD and out of these 34, 5 (14.7%) said there is no difference, 25 (73.5%) said there is improved acceptance and 4 (11.8%) said there is highly improved acceptance. This shows that there is a wide margin between improved and no difference. And putting together the value of improved and highly improved which gives the value of 29 (85.3%) further widen the gap. This shows that MTD acceptance has improved tremendously in the last 2years. The value for no difference, improved and highly improved are the same in Table 5.44 and 5.45. This is a confirmation to the value we earlier have in Table 5.44.

5.7 Summary

This Chapter has presented data analysis carried out in this research with the quantitative data collected. This was done in two categories which include the focus group of UK Oracle Users group and analysis of experts' responses from the online questionnaire administered across the world.

The focus group data were analysed using a quantitative statistical method called Weighted Score or Numerical Indicator. While the data of the experts were coded into SPSS, Percentage Frequency Distribution and Cross Tabulation tools were used. This Chapter also presented the results of the experts' data analysis using a predictive analytical method called Relative Importance Index (RII) to determine the degree of impact of each factor on the adoption of MTD. The use of percentage frequency distribution, relative impact index and cross-tabulation methods of analysis have help in identifying some findings. The discussion on these finding has also led to some conclusions. Having carried out the analysis of the data and the discussion on the findings, the next Chapter will present the development of guidelines from these findings.

CHAPTER SIX

6.0 Development of Guidelines.

This Chapter presents the development of the guidelines based on the findings from Chapter Five. More details on the rationale behind the formation of the guideline will be explained. The guidelines developed will help in the formation of a framework. Furthermore, the framework will be compared to the initial one in Figure 5.1 of Chapter five to highlights the similarities and differences. Insight from the survey has assisted in providing a robust view on issues related to the subject under consideration.

6.1 Development of Guidelines

This section presents the guidelines for an intending subscriber of MTD based on the findings from this study. These guidelines will subsequently be developed into an expert system in the next Chapter. These guidelines are as follows:

G1 – **Experts in the field of Database management system should be given the responsibility of heading the project team when considering MTD adoption**. This is in support of Bajaj (2000) research findings which put forward that a very good knowledge of the new innovation by the management and the project team is very influential to the adoption of the technology. They can be able to present the new innovation and answer any questions related to it confidently with confidence and enthusiasm.

G2 – **The economic factor must be considered which includes the cost and time factor as explained in earlier sections.** This is support of the results from the analysis which shows that the economic factor has the greatest impact on the adoption of MTD. This also supports the findings from Gangwar (2015) and Yeboah-Boateny and Essandon (2014) that MTD gives the advantage of both cost and time benefits compared to on-premises database.

G3 – **The level of security you want your MTD to achieve and tolerate should be examining which include data isolation, scalability, flexibility and customization.** The results from the analysis show that security second highest impact on the adoption of MTD but which influences towards rejection of the concept. This is a proof to the propositions made by Kshetri (2013) and Avram (2014) that the major concerns of subscribers to MTD is the issue of security which covers areas like data isolation, data integrity and scalability.

G4 – **The growth rate of the MTD should also be considered in terms of the size of tenant database, number of tenants, number of users per tenant, growth rate of tenants and growth rate of tenant database.**

G5 – **The harmonious balance between the regulations governing prospective tenants should be looked into by both the tenant and the service provider.** There is variability in terms of where the physical data resides, where processing takes place, and from where the data is accessed. Given this variability, different privacy rules and regulations may apply. Because of these varying rules and regulations, by definition politics becomes an element in the adoption of Database as a Service, which is effectively multijurisdictional (Avram 2014). This aspect of administration, restoration and database reorganisation will require a protecting laws or regulations that will safe guard the interest of both the tenant and the service provider.

G6 – **The choice of the MTD model must be determined by evaluating the features of the three model approaches of the concept which includes share machine, share process**

and share table. Among these four different possibilities for the implementation of multi-tenant database systems, the main problem however remains finding the best balance between scalability, isolation, performance and resource pooling; these are issues which also have to be taken into consideration (Luo et al 2015). There is need to draw the balance between these features depending on the area of interest of the organisation.

6.2 The Amendment of Framework

Based on the analysis and the discussion sections above, the hypotheses in Chapter Five are now further proofed to have the following as the final ones. These will now form the amended framework.

Hypothesis 1 – Economic factor drives towards MTD adoption.

The economic factor still shows that the drive towards MTD will lead to adoption following the results of the analysis.

Hypothesis 2 – Security factor drives towards MTD rejection.

The security factor still shows that the drive towards MTD will lead to rejection as a result of the analysis.

Hypothesis 3 – Growth factor drives towards MTD adoption

The growth factor still shows that the drive towards MTD will lead to adoption as a result of the analysis.

Hypothesis 4 – Regulation factor drives towards MTD adoption

The regulation factor does not tend towards both directions as indicated initially in Chapter Five, after the analysis of the responses from the participants it shows that regulation only lead to adoption. Putting all these hypotheses together will form the new amended framework shown below.

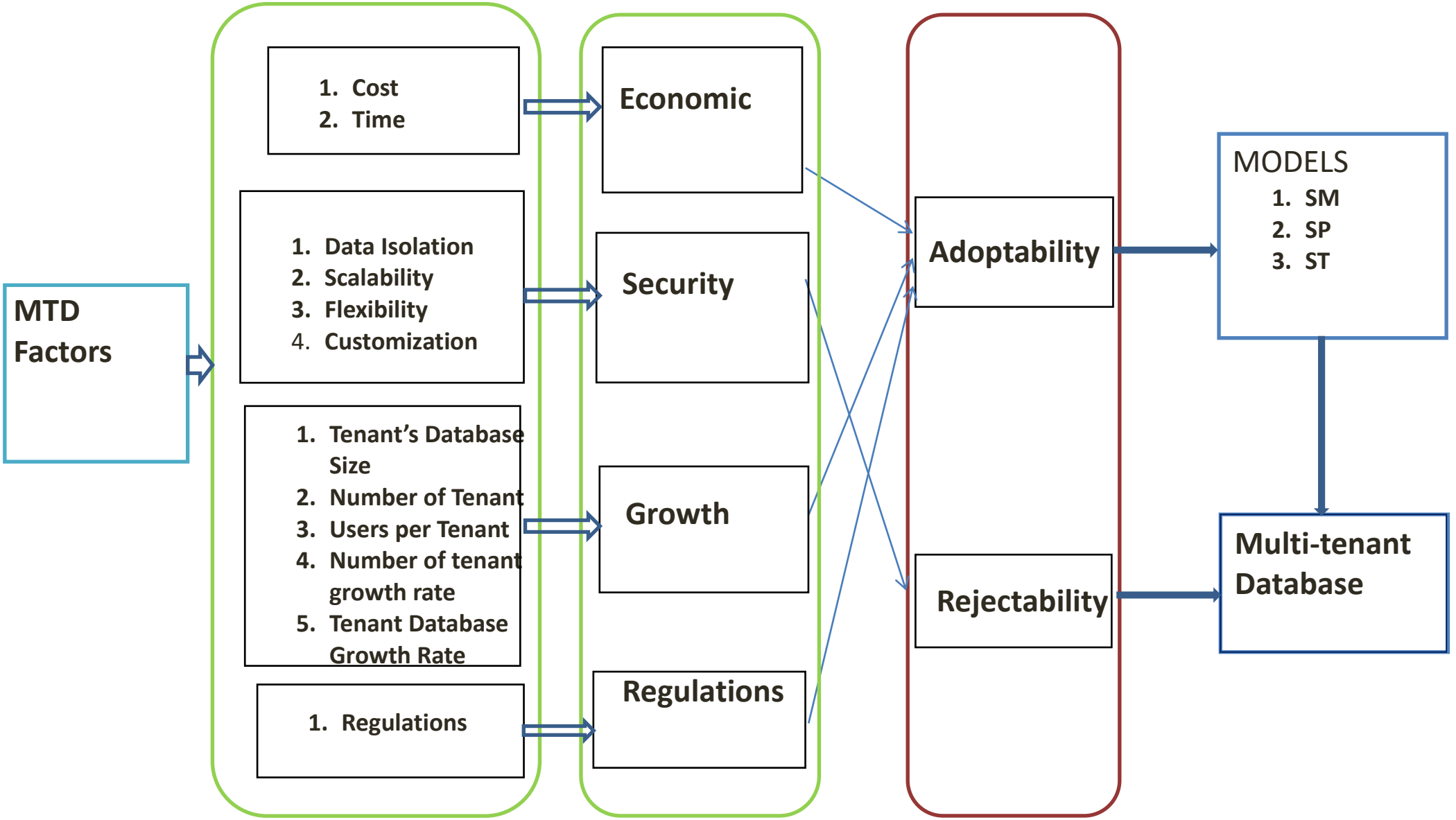


Fig 6.1 - The framework including MTD model type

6.3 The results of the combination of two or more factors

When an intending subscriber wants to know the evaluation impact of considering two or more factors together, this section takes into consideration these types of scenarios. Using the cross tabulation results of MTD usage and the four factors in Tables 6.40 – 6.43 of Chapter Six, the percentage of adoption and rejection for the four factors can be determined as seen in Table 6.1 below. This will also be used in the case of combining of two or more factors by calculating the average cumulative percentage of adoption and rejection for any combined factors. In the same way, using the results from the RII Table 6.27, the RII of any combination of the four factors will be determined by the average of RII of those factors that make up the combination. This above conclusion made about the RII is in line with Gunduz et al (2015) where the factors were rated equally and the results of RIIs for individual delay were used to calculate the average of RIIs of the causes in each group to give the RIIs of the mean group, hence the results shown in the table below. In Table 6.1, E represent economic, S represent security, G represent growth and R represent regulation and others are combinations of two or more of these factors.

Table 6.1 – The Table of Results for Two or Combinations of the Factors

Factor	% Adoption	% Rejection	RII	Final Decision
E	85.7	14.3	0.92	**Adopt**
S	44.1	55.9	0.82	**Reject**
G	76.5	23.5	0.62	**Adopt**
R	76.5	23.5	0.54	**Adopt**
ES	64.9	35.1	0.87	**Adopt**
EG	81.1	18.9	0.77	**Adopt**
ER	81.1	18.9	0.73	**Adopt**
SG	60.3	39.7	0.72	**Adopt**
SR	60.3	39.7	0.68	**Adopt**
GR	76.5	23.5	0.58	**Adopt**
ESG	68.8	31.2	0.79	**Adopt**
EGR	79.6	20.4	0.69	**Adopt**
SGR	65.7	34.3	0.66	**Adopt**
ESR	68.8	31.2	0.76	**Adopt**
ESGR	70.7	29.3	0.73	**Adopt**

In Table 6.1 above the final direction of any of the combinations still amount in the adoption direction, because none of the final percent of rejection is up to the 55.9% as in the case of security that leads to the rejection direction. This means that for the decision direction to be rejection the percentage rejection must be 55.9% or more. Based on this the framework will now incorporate all these possible combinations of the factors. This is now shown below in Fig 6.2.

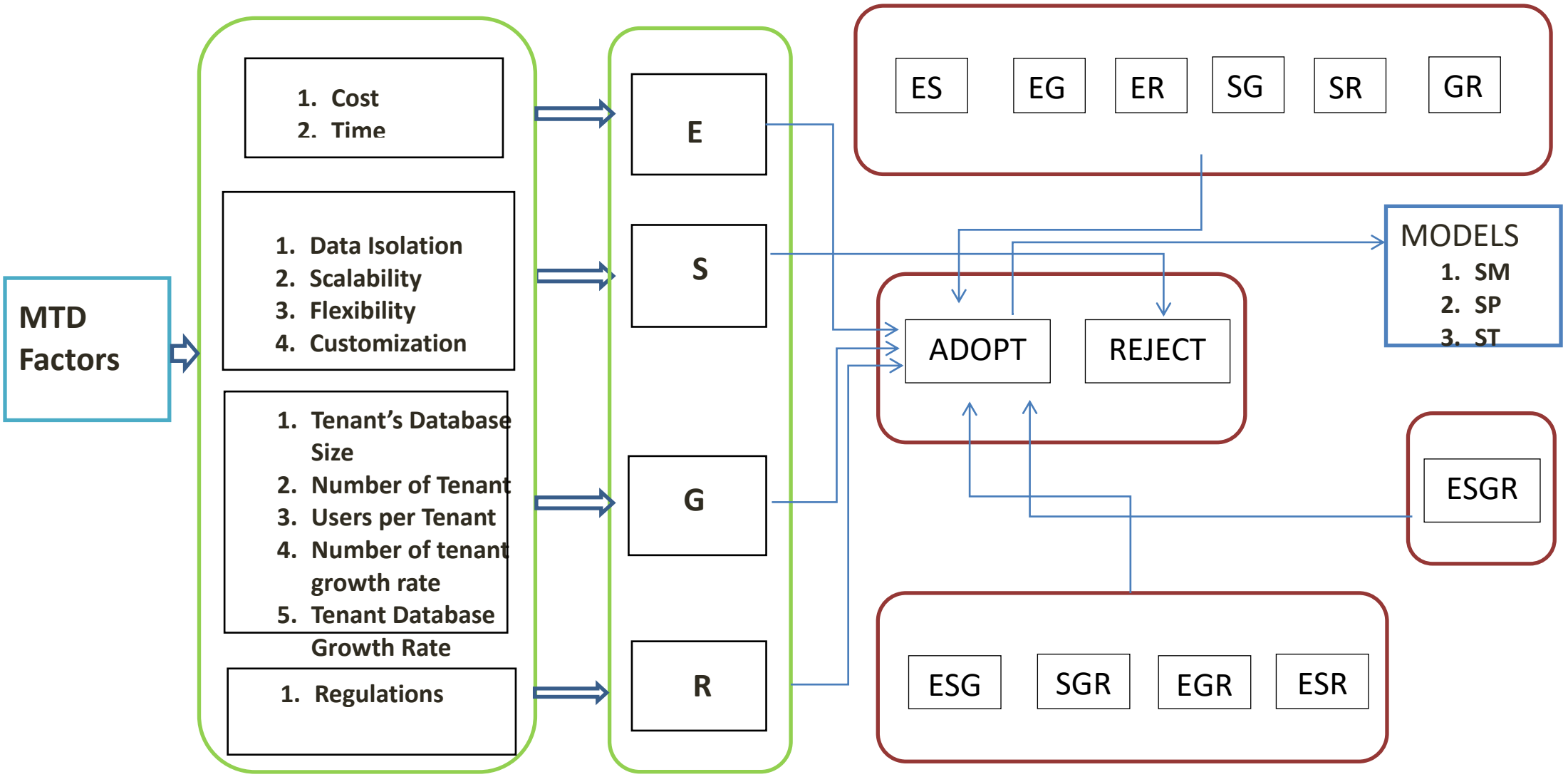


Fig 6.2 - The Modified Framework

6.4 The Comparison of the Initial and New Frameworks.

There are some differences between the initial framework and the new framework. These differences are the results of the analysis of the data collected and discussion from the findings. Looking at the Figure 4.1 and 6.1 there are obvious similarities and also differences.

The similarities are more than the differences, which include economic factors tend towards adoptability, security factors tends towards rejection and growth factors also tend towards adoption.

In terms of differences, we have the regulations as a factor now only tends towards adoption unlike in the initial where it was pointing to both directions. The new framework also considers the MTD model to be adopted considering their different features of each type of MTD models. These models include share Machine (SM), Share Process (SP) and Share Table (ST) as explained in the chapter two.

In Figure 6.2 the modified framework now incorporates the combinations of two or more factors and where these could lead to. This is more comprehensive and user friendly when an intending user wants the benefits of the combination of two or more factor together to inform his or her decision about MTD. This shows that all the possible combinations lead to adoption of MTD. This also means that only the security factor can results in rejection and the impact of the security factor is not strong enough to negate the positive drive towards MTD even when combined with another factor.

6.5 Summary

This Chapter has presented that economic, growth and regulation factors are factors that encourage the adoption of MTD while security factors do not. This Chapter has also shown that the economic factor has the highest degree of impact followed by security and after security is growth while regulation has the least impact. The cumulative impact of two or more factors was also looked into, and the possible directions of these possible combinations were all identified in this chapter. A modified new framework was developed to incorporate the combinations of factors and their directions. Finally a comparison of the initial and new framework was also shown. The next Chapter will show how the framework is developed into an expert system (ES). The method, the tool and the interface the ES will also be illustrated in the next Chapter.

CHAPTER SEVEN

7.0 Development of Framework into Expert System.

7.1 Introduction

In this Chapter, the proposed framework derived is implemented into an expert system which put into consideration all these factors that could be involved in taking the informed decision about MTD by the tenants.

7.2 What is Expert System?

An Expert System (ES) is a computer program that represents and reasons with knowledge of some specialist subject with a view to solving problems or giving advice (Jackson 1999). To solve expert-level problems, expert systems will need efficient access to a substantial domain knowledge base, and a reasoning mechanism to apply the knowledge to the problems they are given and they will also need to be able to explain, to the users who rely on them, how they have reached their decisions (Sun et al 2011). Expert system is a branch of Artificial Intelligence (AI) that makes extensive use of specialized knowledge to solve problems at the level of a human expert (Giarratano and Riley 1998). An expert is a person who has expertise in a certain area. That is, the expert has knowledge or special skills that are not known or available to most people. Expert systems represent a programming methodology by which a computer can be instructed to perform tasks which were previously been considered to require the intelligence of a human expert (Sun et al 2011). An expert system is a computer program designed to imitate a human expert, mimicking the knowledge base and the decision making process of a human expert. An E.S is different from a conventional program because it can explain its behaviour to the human expert and receive new information without new programming.

An expert system is a computer system with the capability of performing at the level of human Experts in some particular domain. It is possible to build expert systems that perform at remarkable Levels (Maher 1987). While there are several methods for designing expert systems, rule-based systems have emerged as the popular architecture. Deriving their knowledge from relatively easily understood Facts and rules, rule-based systems offer surprising power and versatility. Any knowledge based system (referred to as an expert system) essentially emulates the acquired knowledge and thought processes of an expert in arriving at decisions and/or solutions concerning a problem.

Maher (1987) explained that expert systems or knowledge – based expert systems, are interactive computer programs with built in judgement, experience, rules of thumb, intuition, and other expertise to provide knowledgeable advice and solutions on different subjects. Minkarah and Ahmad (1989) provide us with a more specific definition of expert systems as a computer program that uses expert knowledge to reach a level of performance akin to that achievable by highly skilled experts". This is supported by Ye and Wu (2014) that Expert systems are software systems that imitate the decision-making ability of human experts. It is observed that a main distinction of experts and novices in a specialty field is experts' possection of vast amounts of heuristic knowledge acquired and accumulated over many years of experience in the specialty field. Therefore expert systems are designed to address complex problems and to explain the reasoning process, in which the knowledge is represented symbolically rather than numerically. Wijesundera and Harris (1986) describe further the implementation of an expert system as a simulation for a consultation process between an expert of a particular field and a non-expert. Typically, the non - expert is the end user and the computer model is the expert.

7.3 Components of an Expert System

There are two main components of an expert system (Ye and Wu 2014) which are the knowledge base and the inference engine, which performs knowledge-based reasoning to make decisions. During knowledge-based reasoning, the expert system uses a working memory to keep given or inferred facts. Knowledge in the knowledge base can be directly acquired from human experts or extracted through mining data.

The third component of an expert was identified by Sun et al (2011) as the Expert System Interface which the part of the system that interacts with the users of the system. All these three components are further explained in the subsequent sections of this chapter.

7.3.1 The Knowledge Base

This is where the information is stored in the expert system in the form of facts and rules (basically a series of IF statements). This part of the ES has a structure of rules in the form of IF condition THEN consequence, which is also called "Rule Base" (Ye and Wu 2014). This means that when the IF condition(s) are satisfied THEN the consequence will take place. This is where the programmer writes the code for the expert system. This contains necessary

information to solve the problem and this information is obtained from human experts. This is a collection of heuristics which are represented in some manner in the knowledge base.

According to Castillo et al (2012), there are two types of knowledge in an ES: Concrete knowledge and Abstract knowledge. The concrete knowledge is the evidence or facts that are known or given in a particular situation. This type of knowledge is dynamic, that is, it can change from one application to another. These are stored in the working memory which is not permanent in nature. Castillo et al (2012) also explain that abstract knowledge consists of a set of objects and set of rules that governs the relationships among the objects. These are stored in the knowledge base. This type of knowledge is static and permanent, it does not change.

7.3.2 Inference Engine

This applies the facts to the rules and determines the questions to be asked of the user in the user interface and in which order to ask them. This is the 'invisible' part of the expert system, which is active during a consultation of the system (when the user chooses to run the program). Castillo et al (2012) expatiate that the inference engine is the heart of the every ES with the main purpose of drawing conclusions by applying the abstract knowledge to the concrete knowledge. While Ye and Wu (2014) explains inference engine as the aspect of ES that applies knowledge in the rule base to facts in the working memory and make inferences for the goal of making a decision.

The conclusions drawn by the inference engine can be based on either deterministic knowledge or probabilistic knowledge (Castillo et al 2012). The dealing of inference engine with uncertain situations (probabilistic) is more difficult than dealing with certain situations (deterministic). Castillo et al (2012) identifies that one of the weakest elements of ES today is the ability of inference engine to draw conclusions under uncertainty.

An expert system can use 2 different methods of inferencing - Forward Chaining and Backward Chaining.

7.3.2.1 A Backward Chaining system

Backward Chaining is also called goal directed reasoning (Ye and Wu 2014). Works with the system assuming a hypothesis of what the likely outcome will be, and the system then works backwards to collect the evidence that would support this conclusion. Expert systems used for

planning often use backward chaining. This is a top- down approach in which rules are chained together so that the action parts of subsequent rules provide information concerning the validity of the condition part of the preceding rule.

7.3.2.2 A Forward Chaining expert system

Ye and Wu (2014) called forward chaining a data driven system reasoning, which simply means gathering facts (like a detective at the scene of a crime) until enough evidence is collected that points to an outcome. This is the reasoning from the facts to the conclusions resulting from those facts. Forward chaining is often used in expert systems for diagnosis, advice and classification, although the size and complexity of the system can play a part in deciding which method of inference to use. Here the condition part in each rule is checked against the database to establish the validity.

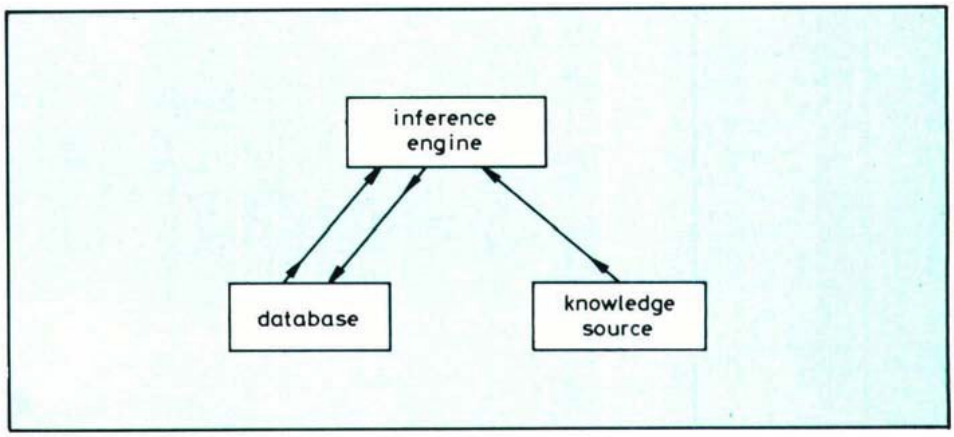


Figure 7.1 Expert System Components (Jackson 1999)

7.3.3 Shell – User Interface.

This is where the user interacts with the expert system. Castillo et al (2012) defines the user interface as the liaison between the ES and the user, while Giarratano and Riley (1998) define as the mechanism by which the user and the ES communicate. The incorporation of efficient mechanisms to display and retrieve information in an easy way makes ES an effective tool. In other words where questions are asked, and advice is produced. As well as the advice that is output, the user interface can output the justification features of an expert system. Examples of information to be displayed are the conclusions drawn by the inference engine, the reasons for such conclusions, and an explanation for the actions taken. When there are no conclusions reached by the inference engine, the user interface provide a vehicle for obtaining more

information needed from the user that will further help the inference engine to get a conclusion. A good and effective ES must provide avenue for this through the interface otherwise the quality of the ES will be in doubt (Castillo et al 2012).

7.4 Development of the Expert System.

This section examines the Expert System shell used in this research to implement the framework, called ES BUILDER. The choice of this shell is largely based on its free access and use of the software. It is also a web based expert system shell. This shell is explained in detail in the subsequent sections below.

7.4.1 ES-Builder

ES-Builder is an Expert System Shell application. The software is used to design expert systems that may be accessed dynamically as web pages and incorporated as a knowledge base in any web site. ES-Builder features a decision tree modelling process for developing the logic of the expert system (ES). The ES-Builder program was built in order to assist expert system developers by providing a simple interface for implementing model expert systems that may have been pre-designed using a suitable design process. This type of expert system is developed using a process of deductive reasoning. This means that the expert system provides an interface to test a series of attributes, which through the process of deduction allows the user to arrive at a conclusion. This conclusion is logically correct based on the values chosen by the user for every attributes involved.

Building an expert system with ES-Builder is easy, because it uses a simple web interface which can be easily accessed by anyone familiar with the internet. The user constructs the expert system using a decision tree interface where attributes, values and conclusions are added as leaf nodes on the tree. Each node has a small integrated data set which is used to form the content of the expert system when it is accessed online. When the expert system is completed and make available on the internet, the user simply has to click on an option from a list presented on a page for each attribute. Attributes are displayed in sequence with only values appropriate to the current search shown.

7.4.2 Using the ES-Builder Expert System Shell

ES-Builder is a dual purpose program. It allows the user to both create an expert system (ES) and you can at the same time search the expert system for conclusions. This allows the user the flexibility to both edit the expert system and to test it within the one application.
Before expert systems can be created in the ES-Builder expert system shell, the user first needs to carry out the following steps:

- Plan and design the expert system efficiently using a clearly defined Universe of Discourse (UofD)
- Have identified all the conclusions to be included in the expert system
- Have determined all of the attributes that will be tested by the expert system
- Have researched the UofD thoroughly and have identified all appropriate values for each attribute.

In addition, the expert system can be improved by:

- The preparation of extra notes on each possible attribute, value, and conclusion to further inform the user about the result of each search
- The use of suitable image graphics to illustrate each attribute, value, and conclusion during the search process, this is optional.

7.4.3 Searching the Expert System

The process of searching on ES built by ES-Builder is very simple. Starting from the first hyperlink on the home page of the published expert system, the user can click on the most appropriate response to the attribute tested on each page. The system moves to the next attribute or reaches a conclusion. The user may use the back button in the browser at any time, if a wrong choice has been made. These steps can be repeated until all the desired attributes are taken and a conclusion is reached.

7.4.4 Understanding the Search results

The search result gives the conclusion from the expert system that matches the responses the user has made. This process of deduction assumes that the user has made accurate observations and given accurate responses.

This type of expert system has no Artificial Intelligence and cannot make an educated guess at what it is the user may be observing. This form of knowledge engineering relies completely on the rules created by the knowledge engineer that is the ES developer during the development stage. The rule that results in a conclusion being found is reported on the Search Results page. Any additional information recorded in the expert system along with this conclusion is also displayed in this page.

7.4.5 Using ES-Builder Web

Creation of an expert system (ES) in ES-Builder is only possible for registered users of ES-Builder Web. Firstly, this involves creating a user account via the ES-Builder Web User Registration page. Each user must supply a unique email address for registration as a username. User email addresses and passwords are stored in the database. This registration is confirmed by the supplied email address before a user may login in to the system.

7.4.6 Building an Expert System

Firstly, users must create a new project for their expert system design. This is possible once logged in.

7.4.6.1 Project Details

In the Project Details page, the user defines a title for their expert system, defines the Universe of Discourse, may define an image to display on the title page, and edit other settings for the expert system. All these are shown in the Figure 7.2 below

Figure 7.2 – MTD Project Details Page

7.4.6.2 Decision Tree

The deductive logic of the ES is created through the Decision Tree View by entering the title details, attributes, values, and conclusions into a decision tree. Each step in the decision tree is called a node. A node that branches out of another node in the decision tree is called a branch node. A node may have branches to further nodes, and so on, until the decision tree is complete. There are a number of basic rules about how the tree can be formed and which branch nodes a particular type of node may accept. The nodes at the very ends of branches are called leaf nodes. The decision tree for this MTD evaluation framework is shown in the Figure 7.3 below.

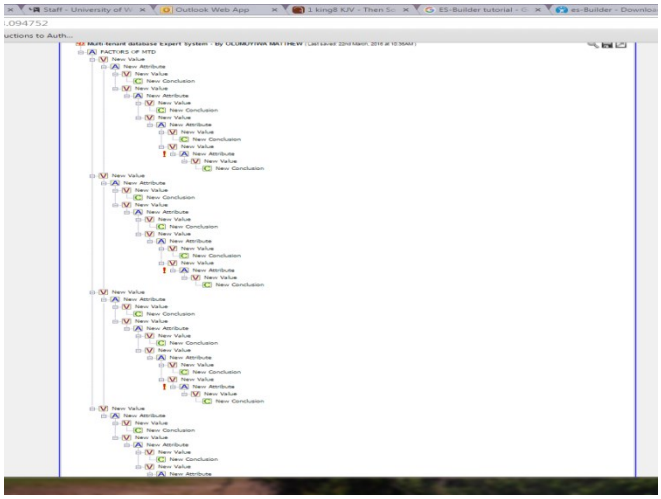


Figure 7.3 – MTD Decision Tree

7.4.6.3 Universe of Discourse

The first (or root) node in any decision tree is the Universe of Discourse (UofD). Details about UOfD are entered by the user when they create the project including:

- The name of the ES.

- The identifier to be used to refer to each conclusion in the ES.
- A phrase to be used as a starting link at the beginning of the ES.
- A longer description of the ES to be presented on the home page of the ES. This longer description can be created using HTML tags to improve presentation in the browser.
- An image to be displayed on the home page of the ES to improve presentation.

7.4.6.4 Attribute

The second node (or first branch) of the decision tree must be an Attribute which is displayed with an 'A' icon. Attributes are characteristics of possible conclusions that are to be tested in the ES. Each Attribute must have at least two branch nodes in the completed system. The only type of branch node accepted by Attribute nodes are Value nodes.

7.4.6.5 Value

Each Value node represents the most correct response to an Attribute for a particular conclusion. Value nodes may have two possible types of branch node: Attribute and Conclusion. When a further Attribute needs to be tested, the branch node of a Value will be another Attribute node. When a final conclusion has been made, the branch node of the Value will be a Conclusion node. Value nodes may have only one branch node.

7.4.6.6 Conclusion

A Conclusion node must be a leaf node. No branches are accepted from Conclusion nodes.
For each node apart from the first (UofD) node, three data items can be added. Each Attribute, Value, or Conclusion node may have:

- a detailed definition (this allows the designer to use a short identifier in the tree to keep the design process neat and simple)
- a paragraph of help notes to inform users about the process of the ES and to give more detailed information about possible conclusions
- An image to be displayed in the ES to assist users in the process and to give detailed visual information about possible conclusions.

7.4.6.7 Knowledge base

The knowledge base is captured in the Figure 8.4 below showing information stored in the ES in the form of facts and rules. This part of the ES has a structure of IF condition(s) THEN consequences. This means IF statements are satisfied the THEN will take place.

1	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? ECONOMICS **AND** ECONS ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
2	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? ECONOMICS **AND** ECONS ONLY NO **AND** AND SECURITY ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
3	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? ECONOMICS **AND** ECONS ONLY NO **AND** AND SECURITY ONLY NO **AND** AND GROWTH ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
4	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? ECONOMICS **AND** ECONS ONLY NO **AND** AND SECURITY ONLY NO **AND** AND GROWTH ONLY NO **AND** AND REGULATIONS ALSO YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
5	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? SECURITY **AND** SECURITY ONLY YES **THEN** the MTD Adoption or Rejection is **REJECT MTD**.
6	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? SECURITY **AND** SECURITY ONLY NO **AND** AND ECONS ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
7	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? SECURITY **AND** SECURITY ONLY NO **AND** AND ECONS ONLY NO **AND** AND GROWTH ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
8	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? SECURITY **AND** SECURITY ONLY NO **AND** AND ECONS ONLY NO **AND** AND GROWTH ONLY NO **AND** AND REGULATIONS ALSO YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
9	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? GROWTH **AND** GROWTH ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
10	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? GROWTH **AND** GROWTH ONLY NO **AND** AND ECONS ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
11	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? GROWTH **AND** GROWTH ONLY NO **AND** AND ECONS ONLY NO **AND** AND SECURITY ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
12	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? GROWTH **AND** GROWTH ONLY NO **AND** AND ECONS ONLY NO **AND** AND SECURITY ONLY NO **AND** AND REGULATIONS ALSO YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
13	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? REGULATIONS **AND** REGULATIONS ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.
14	**IF** WHICH OF THESE FACTORS IS MOST IMPORTANT TO YOU? REGULATIONS **AND** REGULATIONS ONLY NO **AND** AND ECONS ONLY YES **THEN** the MTD Adoption or Rejection is **ADOPT MTD**.

Figure 7.4 – MTD Knowledge base

7.4 Summary

There is a need for public sector organisations and general IT investors to embrace MTD platform as means of securing data because of the cost saving associated with it compared to investing in a dedicated database managements system (DBMS). This Chapter has presented an Expert System (ES) incorporating the proposed framework derived in the earlier chapter. The expert system was developed using the web based development tool ES-BUILDER (Mcgoo.com 2013) for easy access and use. All the stages involved in the ES development were presented in this chapter. For further understanding, the link to the ES is as follows

http://www.mcgoo.com.au/esbuilder/viewer/viewES.php?es=252e59e9368580a68e0b52630b4c6f27

The next Chapter will present the validation process of this research. It will show the analysis of the responses from the validation method and the results of the validation.

CHAPTER EIGHT

RESEARCH VALIDATION

8.0 INTRODUCTION

This Chapter discusses the validation of the research guidelines and framework developed in the research programme. The framework subsequently led to an expert system being developed which will serve as a guide to assist intending users of MTD to take an informed decision about the adoption of MTD. This will create an enabling environment and enhancing the capacity of the intending MTD users during the process of evaluating and adopting the concept.

Having developed a framework and the expert system, there is a need to test its validity before it can be more widely disseminated or used. The aim of the validation process is to determine whether the research findings and recommendations are reliable. The reliability and objectivity of the research is revealed through the validation process. It is important to carry out validation not only to prove the genuineness of the findings but also to look at all stages involved in carrying out the research. It is essential for any scientific inquiry that researchers ensure the quality of their work in every step of its methodology, including data collection, analysis, and interpretation of results, through appropriate validation techniques (Lucko and Rojas 2010). The results and the process by which they were derived need to be accepted by the academic and the professional communities, so that the new knowledge becomes another stepping stone in the advancement of the state-of-the-art and filters down to daily practice, ultimately contributing to the welfare of society. Validation provides a solid background against which the research findings could be generalised. The next section provides a general discussion of the concept of validation then the method adopted for undertaking the validation exercise.

8.1 The Concept of Validation

The purpose of validation is to ensure the integrity of all techniques and procedures used for the development of theorem in a research in order to establish confidence in those techniques and procedures. Validation is a key part of the model/framework development process which increases confidence in the model/framework and makes it more valuable (Kennedy et al., 2005). Validation is the process used by the scientific community to acquire the necessary information to assess the ability of a technique or procedure to reliably obtain a desired result.

Validity determines whether the research truly measures that which it was intended to measure or how truthful the research results are (Golafshani, 2003).

8.2 Framework and Expert System Validation.

The proposed Expert system has to be validated in order to confirm whether it is appropriate especially in the light of the contest of the purpose of this investigation (Frees, 1996). Egbu (2007) describes the validation of a model/framework as the process of assessing the ability of the model/framework to do what it sets out to achieve. This process attempts to ensure that the model/framework represents the characteristics of the general population and is not peculiar to the samples used in its estimation (Hair et al., 1998). According to Ankrah (2007), the validation process thus seeks to assess the extent to which the models predict the outcomes in terms of performance above or below average. There are two components of validation namely; Internal and External validation.

Egbu (2007) notes that internal validation seeks to outline the strength of the model/framework as well as assess the literature search. Prior to implementing a new model or technique, the reliability of the technique must be demonstrated which include testing the technique using similar samples and conditions, if a modification has been made, the modification must be compared to the original technique. This is what internal validation does where consistency and reproducibility must be determined by repetitive analyses.

External validation must have been carried out by a scientific, scholastic, or professional body other than that of the immediate organisation prior to the implementation of the new technique or model. Brinberg and McGrath (1985) state that the essence of external validation is to gain confidence in the findings and what they mean. It is about ensuring the robustness of the research and about assessing its generalisability (Rosenthal and Rosnow, 1991; Fellows and Liu, 1997). Internal validity concerns the credibility of the inferences made from the data while external validity concerns the generalisability of the findings (Eisenhardt and Howe, 1992; Kirk and Miller, 1986). All validation procedure must also be documented to ensure that any qualified individual could evaluate what was done and replicate the validation process.

This research adopted both internal and external validation methods. Having developed the guidelines, framework and the expert system showing how to evaluate and adopt a Multi-tenant database, there is need to test the validity of the results from the perspective of the intending users.

8.2.1 Selection of Participants for Validation

There are three options that can be considered for carrying out the validation, which include focus group, interview and online surveys. Online surveys are constrained by their restrictive nature and lack opportunity to clarify respondents' unclear views was handled by carefully designing the questionnaire. The online survey was now considered over focus group. Also, few interviews were conducted because of time and cost constraints. A copy of the research framework and a link to the web-based expert system was attached to the survey to clarify any misunderstandings the respondents may have.

It was important to validate the findings with stakeholder in the market of DaaS, to determine if the findings were valid and the recommendations useful in respect to their experiences. Based on this reason, a covering letter was sent via email to the participants that were initially involved at the early stage of the research including participants who participated in the focus group and the online participants. The use of the previous participants is based on their prior involvement in the earlier survey, which makes them familiar with the research and possibly ensures a good response rate. Taking one's findings back to the subjects being studied where the people can verify the findings, has been argued by Silverman (2006) as being that one can be more confident of their validity. This method is known as respondent validation (Silverman, 2006).

Also, validation of the framework helps to ensure that the research has actually identified key factors affecting MTD adoption amongst organisations in the public sector and has sought to assess the extent to which the framework endeavours to enable intending users to make an informed decision about its adoption. That is, if the framework has provided accurate steps to take in evaluating and accessing the concept with respect to its adoption and effective use by both users and providers. The next section therefore describes the validation process and the conclusions drawn from the findings. This would also help to predict if the usefulness of the research outcome was about, or above average. The following sections provide a description of the methods adopted for the validation exercise.

8.3 Methods Adopted for Validation

There are two methods adopted for the validation process of this research, which includes external and internal validation.

8.3.1 External Validation

External validation aims to address the accuracy of a model in a domain using a different but plausibly related population, which may be defined as a selected study population representing the underlying domain (Bleeker et al 2003). Yin (2013) describes external validity as determining the limits to which the findings of the research could be generalised. Brinberg and McGrath (1985) state that the essence of external validation is to gain confidence in the findings and what they mean. In other words, it is the extent to which the results of a study can be generalized to other situations and to other people. It is about ensuring the robustness of the research and about assessing its generalisability (Rosenthal and Rosnow, 1991). External validity is the degree to which the conclusions in your study would hold for other persons in other places and at other times.

External validity was achieved in this research by comparing the findings with similar findings from previous studies (Eisenhardt, 1989). Participants who took part in the first and second phases of the research were invited to share their opinions on the research findings and recommendations in a questionnaire survey. Although the sample size used for this validation exercise is relatively small, the feedback received is generally encouraging and suggests that the research findings and recommendations have the potential of being well received. The outcomes suggest that the findings and recommendations are useful in terms of stimulating the adoption of MTD. The feedback also creates assurance that the developed framework could assist the intending users, service providers as well as other stakeholders in increasing the adoption and effective utilisation of MTD in the public sectors. The tables below present a summary of the results that were obtained from the participants who responded to the questionnaire. Indeed, results from the questionnaire and some of the positive recommendations made by a number of the participants acknowledged that the framework is useful and would serve as a detailed guide for the major groups that are involved with MTD adoption and usage.

8.3.2 Participants Response

The validation survey was conducted online using Google Form (MTD Validation Survey 2016). 21 responses were received. Out of the 21, 16 were made up of database experts that participated in the initial survey, while 5 were other database experts from different organisations other than the original 16. The interview stage comprises of 4 experts, one from the category of those that participated in the initial survey where the organisation of this

participant is in use of MTD while the second interviewee is not part of those that participated in the initial survey but works for a MTD user organisation.

The data was analysed using SPSS to determine the frequency and percentage to which respondents at least agree to the research outcome (see appendix F). The majority of the participants were in favour of the outcome indicating that the framework is capable of assisting individuals and organisations in taking an informed decision about the adoption of MTD. The responses also show that the majority agree with the outcome that the framework has incorporated all aspects needed for the decision making process. Finally, the majority also agree that the expert system is simple and user friendly enough for the intending tenant(s) to support their process of making decision in regards to MTD. All the results received were to a large extent positive as shown in Tables 8.1, 8.2 and 8.3 below.

Table 8.1 – Validation of research findings (Frequency-Percentage)

Item (Question)	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Cost	0 (0%)	0 (0%)	0 (0%)	7(33.3%)	14(66.7%)
Time	0 (0%)	0 (0%)	3(14.3%)	9(42.9%)	9(42.9%)
Economic	0 (0%)	0 (0%)	0 (0%)	7(33.3%)	14(66.7%)
Economic Impact	0 (0%)	0 (0%)	0 (0%)	12(57.1%)	9(42.9%)
Growth	0 (0%)	0 (0%)	3(14.3%)	9(42.9%)	9(42.9%)
Growth Impact	0 (0%)	1(4.8%)	6(28.6%)	5(23.8%)	9(42.9%)
Security	0 (0%)	0 (0%)	0 (0%)	9(42.9%)	12(57.1%)
Security Impact	0 (0%)	1(4.8%)	0 (0%)	9(42.9%)	11(52.4%)
Regulation	0 (0%)	1(4.8%)	3(14.3%)	7(33.3%)	10(47.6%)
Regulation Impact	0 (0%)	0 (0%)	2(9.5%)	11(52.4%)	8(38.1%)

Table 8.2 – Validation of research recommendations (a)

Item (Question)	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Experts	0 (0%)	0 (0%)	1(4.8%)	12(57.1%)	8(38.1%)
Economics Consideration	0 (0%)	0 (0%)	1(4.8%)	5(23.8%)	15(71.4%)
Level of Security	0 (0%)	0 (0%)	0 (0%)	10(47.6%)	11(52.4%)
Growth Rate	1(4.8%)	1(4.8%)	6(28.6%)	8(38.1%)	5(23.8%)
Regulation Balance	0 (0%)	1(4.8%)	2(9.5%)	8(38.1%)	10(47.6%)
MTD Model Choice	0 (0%)	0 (0%)	0 (0%)	9(42.9%)	12(57.1%)
Framework Completeness	0 (0%)	2(9.5%)	2(9.5%)	11(52.4%)	6(28.6%)
ES Friendliness	0 (0%)	0 (0%)	1(4.8%)	9(42.9%)	11(52.4%)

Table 8.3 – Validation of research framework capability

Item (Question)	Not sure of its capability (1)	No, not capable (2)	Neutral (3)	Yes, Capable (4)	Yes, Highly capable (5)
Framework Capability	0 (0%)	0 (0%)	1(4.8%)	13(61.9%)	7(33.3%)

Based on the responses received on the research findings, it shows that the majority of the respondents agree or strongly agree to the findings. The response from all items have more than 80% cumulative percentage score for both agree and strongly agree except in the case of Growth Impact which has a cumulative percentage value for both agree and strongly agree of 66.7%. Therefore, the overall research findings were largely accepted by all the respondents.

The results with respect to the research recommendations are also presented in Table 8.2. The results from recommendations put forward to support intending users and service providers of MTD were all accepted by the respondents. This shows that the findings and recommendations are all valid. In terms of the framework completeness in Table 8.2 which fully incorporated all the aspects needed for this decision making process, there is a cumulative value of 81% for agree and strongly agree. From the same table, the ES Friendliness, which has to do with the simplicity and user friendliness of the ES for the intending tenants, has a cumulative percentage value for agree and strongly agree of 95.3%. In addition, as can be observed from Table 9.3, most of the respondents agree that the

framework is capable of supporting users in taking decision about MTD with a cumulative value of 95.2% for both agree and strongly agree. All these suggest that the research would be regarded as a very useful tool for decision making as more than 65% of the participants' opinions in all items were in favour of the research findings. This represents a positive contribution to the body of knowledge.

Interviews were also conducted with one IT managers, one database administrator and two users of an organisation where MTD is in use to make their own assessment of the research framework and the ES, to also offer suggestions on how the framework or ES could be improved. Some of these respondents re-emphasised that the research findings have high potential of achieving its purpose.

There were a few interesting assessments made which are noted below:

"Your research is well structured and relevant to the adoption of Multi-tenant database by intending individuals or organisations. I couldn't have thought of anything better than your new framework" [Service Provider - Manager].

"*The findings and recommendations show that an in-depth research, consultation and analysis have been done. This will be very useful in the process of adopting Multi-tenant database model"* [Organisation – DBA].

"*Your framework in conjunction with the ES will be very helpful. I hope organisations and tenants consider a number of the recommendations you have put forward in your research and use it as a guide to ease the process of Multi-tenant database adoption*" [IT Officer].

Some of the participants made few suggestions:

"*The framework is very detailed and will be useful but I think maintainability should be incorporated into the framework*" [Database User].

A respondent also notes that:

"*The framework and the ES are very resourceful and should made available to the public for easy access*" [IT Manager].

Based on the above responses from the participants, external validation has been used to substantiate the research findings. Internal validation is also used as explained below.

8.3.3 Internal Validation

Internal validity refers to how well an experiment is done, especially whether it avoids confounding (McDermott 2011). Internal validity refers to how well a piece of research allows you to choose among alternate explanations of something. A research study with high internal validity lets you choose one explanation over another with a lot of confidence, because it avoids (many possible) confounds. Rosenthal and Rosnow (1991) define internal validity as the degree of validity of statements made about whether X causes Y – the primary concern being to rule out plausible rival hypotheses. Egbu (2007) notes that internal validation seeks to outline the strength of the model/framework as well as assess the literature search. Internal validity concerns the credibility of the inferences made from the data while external validity concerns the generalisability of the findings (Eisenhardt and Howe, 1992; Kirk and Miller, 1986). Interestingly, whilst researchers agree that both internal and external validation are important for validating a research process, there are limited literatures that describe in detail what form the internal validation process should take (e.g. Fellows et al., 2002).

This study has adopted some measures to achieve internal validity. The first measure adopted was feeding back responses and findings to the participants as suggested by Easterby-Smith et al (1991). This enabled the participants to check the accuracy of their responses that is the accounts of the participants are factually correct. This also presented an opportunity to the participants to provide feedback to the researcher's interpretation. The feedback has enhanced the study's interpretive validity as argued by Maxwell (1992).

It is also important to note that some findings of this research have been presented and published in a number of international peer reviewed journals and conference proceedings as listed in Table 8.4. And most of the arguments and findings of the research were supported by comprehensive literatures research.

Publication of articles in international academic journals and conference proceeding is a means of disseminating research findings to the academic community. This involves a review and assessment of the validity of research and its finding by independent referees. A total of four articles have been published, which include two journals and two conference articles, with one currently under review. Xiao (2002) states that peer review in this manner provides an opportunity for the methodologies, meanings and interpretation of the research to be questioned. Runeson and Loosemore (1999) refer to this dissemination process as a process

of critical inquiry which is meant, in theory, to provide an informed, fair, reasonable and professional opinion about the merits of the research. Fenn (1997) has observed that peer review is used as the gold-standard throughout academia in the UK. Feedback from such a process helps to enrich research work and potentially improves its findings (Alkass et al., 1998). The feedback provided by referees always shows the reasons for their points and views. All these points raised were incorporated in this study to improve the validity of the research.

The journals targeted so far are International Journal of Information Technology and Computer Science, International Journal of Computer Science and Emerging Technologies. While papers have been presented and published at the 19th UK Academy for Information Systems (UKAIS) Conference (2014), UK and 1st International Conference on Computer and Information Science and Technology (CIST'15), Canada.

The remarks and feedback from the academic community during the presentation and review that have been incorporated in the research and into this thesis, have significantly improved the research, making the findings more robust and reliable, as argued by Xiao (2002). Acceptance of the articles for publication indicates that this research is scholarly and academically valid. However, the dissemination of the research findings with respect to academic validity is on-going with intended papers aimed at focussing on specific and distinct aspects of the research findings. The Table 8.4 below presents a list of journal and conference papers that have been published.

Table 8.4 – Published Journal and Conference Articles

No	Authorship	Year of publication	Type of Publication	Remarks
1	Matthew, Dudley and Moreton	2014	Conference Paper	Published
2	Matthew, Buckley and Garvey	2015	Conference Paper	Published
3	Matthew, Buckley and Garvey	2016a	Journal Paper	Published
4	Matthew, Buckley and Garvey	2016b	Journal Paper	Published
5	Matthew, Buckley, Garvey and Moreton	2016c	Journal Paper	Published

8.4 Summary

This Chapter reports on the validation of the research findings, recommendations, framework and expert system. The Chapter describes the validation process, which includes both external and internal validation. The internal validation was based on academic validation which involved the publication of some aspects of the research findings in journals and conference proceedings. In these papers, a significant number of references have been cited to support the different arguments. Moreover, the concepts, methodology and findings of this research have been found to be reasonably supported by the extensive use of literatures in support of the study. The external validation involves respondents who participated in the empirical data gathering phase who were invited to share their opinions on the findings and recommendations. The views from both areas were reported within this Chapter of the thesis. The results from the analysis of the participants' responses indicate that the findings reported in the research are valid and can be generalised across the world of DaaS. Likewise, the majority of the respondents who shared their opinions with regard to the findings, to a large extent agreed with the findings. In the next Chapter, the conclusions of this research based on the analyses and validation process will be presented. The research limitations and recommendations for further research will also be put forward.

CHAPTER NINE

CONCLUSIONS AND RECOMMENDATIONS

9.0 Introduction

This Chapter aims to give the overall conclusion of the research and is divided into seven sections. This section provides detailed explanation of the entire study which led to the research findings including the factors that influence adoption of MTD, the guidelines, the framework and the expert system for the evaluation and adoption of MTD. The first section provides an overview of the research. Discussions based on the accomplishment of the research aim and research questions are in the third section. The fourth section provides the key contributions of the research in terms of knowledge and practice. The fifth section provides some of the research limitations as observed by the researcher and suggestions for further research are presented in the sixth section.

9.1 An Overview of the Research

Deploying Database as a Service (DaaS) is gaining momentum with a significant increase in the number of organisations ready to take advantage of the technology. This architecture will reduce the cost incurred in the development and deployment of an on-premises database system both in hardware and software required for such a system. The multi-tenant data management system amortizes the cost of hardware, software and professional services to a large number of tenants and thus significantly reduces per- tenant cost by increasing the scale.

There are still some factors that need to be considered before allowing data to be hosted by a third party database service provider's platform. These factors are identified in this research in chapter three of this thesis and what is the degree of impact each one has on the adoption of MTD. However, the findings from this research have provided a comprehensive report on the perspectives of database experts from the public sectors across the entire world. The samples for the research were carefully selected from different continents of the world. The research incorporates elements of previous studies in multi-tenant databases and with empirical data, to address the research concern.

The overall purpose of this research has been to ascertain the factors affecting the adoption of MTD and show the level of impact each factor has, furthermore, the research is aimed at developing a framework to help the intending users in making an informed decision. It is

intended that the recommendations put forward, based on the empirical findings in this research would help to provide a guide to DaaS providers and users.

9.2 Overview of the Research Findings and Outcomes

In this research, a major role was played in the area of investigating the different factors that influence the evaluation and adoption of the multi-tenant database concept. The main objective of the research was to develop a framework that could assist in resolving the challenges faced by intending users of MTD with regards to these factors before adopting any model of the concept. The literature review aspect of the research revealed these factors, but no theoretical perspective consideration has ever been carried out in this area to determine the impact level of these factors.

A thorough and careful study was done in this research which has analysed these factors and conducted a quantitative research survey. A framework was developed and subsequently converted into an expert system. These were further tested and validated with the involvement of experts and professionals in the field of database management. The subsequent sections briefly present and discuss the significant findings of each phase, then examine whether the research aim was achieved. Another section addresses the limitations of the study and some possible future research directions.

9.3 Achievement of Research Aim and Questions

The aim of this research is to develop a standard scientific guideline that will support the public sectors in their quest to make an informed decision towards the evaluation and adoption multi-tenant database approach, putting into consideration several factors that could influence this decision. This aim has been achieved having identified these factors with a thorough literature analysis coupled with two stages of data collections and analysis which has been done in this research. The research has helped determine the impact of each factor on the decision to adopt MTD which had led to the development of the framework.

The first research question considered the different methods used in the implementation of the MTD model. The question was answered as the research has identified three different methods of MTD model in chapter two. They included the share machine, share process and share table. The research found out that each of these methods has its own peculiarity based on their features.

The second research question addressed the level of data privacy and isolation each of the models has. The research found that each of the models has its own level of data isolation they can offer. Share table has the lowest degree of data isolation followed by the share process method while the share machine has the highest level of data isolation. In the effort of finding answers to the second question it was discovered that the share machine has a very poor scalability, the share process has good scalability and the share table has the best and the highest level of scalability. There are other features identified which are summarized in Table 2.2 of Chapter Two where the answers to research question two are presented.

The third research question attempted to determine in details the factors that influence the decision to adopt a Multi-tenant database model. This question was addressed by conducting a thorough and extensive literature reviews from both an academic and professional perspective to achieve this in Chapter Two. These factors include size of tenant database, number of tenants, number of users per tenant, growth rate of tenants, growth rate of tenant database, cost, time, flexibility, regulation, scalability, customization and data isolation. These were further classified into four groups in this research for easy analysis. This grouping includes economic, growth, security and regulation.

The degree of impact or influence each of these factors has towards the drive for MTD was addressed by the fourth research question. This was achieved after the data collection and analysis carried out for both the focus group and the professionals that participated in the surveys. The results obtained after the data analysis assisted in identifying the impact of each factor which was achieved in Chapter Five. This has led to some recommendations being put forward from the results. And a framework was developed that can facilitate a successful adoption of MTD. It will serve the intending users and service providers of MTD to make an informed decision about the adoption of MTD.

The fifth research question identified the acceptance level of MTD in the last two years. Part of the empirical data based on results obtained from data collection were analysed to answer this question in Chapter Five. The research observed that prior to the last two years (2013-2015) there was not much difference in the acceptance of MTD but the research further observed that there is improved level of acceptance to the value of 73.5% in 2016.

The research has also proofed that its findings are also valid for private organisations. The participation of experts was not limited to public sectors, therefore the findings are also viable for private and all types of organisations.

Overall, in addressing the research questions and objectives, the research design, approach and method adopted has proved its genuineity and validity with the findings, recommendations, framework and expert system developed from the research.

9.4 Research Contributions

This research has contributed to the existing body of literature and the field of Information Technology by conducting a thorough review on the concept and implementation models of MTD. This research has empirically identified key factors that influence the adoption and utilisation of Multi-tenant database model.

No previous study had empirically considered the evaluation and adoption process of the types of MTD models. There is lack of scholarly articles on this process amongst the public organisation on the possibility of giving their data to a third party organisation. Therefore, this study adds to the existing body of literature and makes specific contributions to the field of IT by providing insights on these influencing factors about the adoption.

It was observed that no previous research had put forward a guide for resolving the various issues and concerns facing the intending MTD users in the adoption of a model. Hence, this research is considered as one of the pioneer studies in this area. The research has been able to put together a guide in the form of a framework that assist both the intending users and service providers of MTD. In other words, creating a novel framework to aid the successful decision making process, adoption and utilisation of MTD constitutes the central contribution of this research. Overall, the framework provides a dynamic view of all the influencing factors of MTD evaluation and adoption as the recommendation for improving the understanding of the process for a successful MTD adoption.

This research has also made a novel contribution to the area of MTD adoption as it has identified the degree of impact each factor has on the decision about MTD. This will help to determine which of the factors is considered more important to the user depending on the area of usage by the user. This aspect has not been identified in previous research.
The level of MTD acceptance in the last two years (2013-2015) was also one of the contributions made by this research, which has not been identified in previous research.
This research has also made a novel contribution to the area of MTD adoption by developing the framework into a web-based Expert System. This ES has incorporated all aspects of the framework for easy access and use by the intending users.

The research framework developed can be applied by other researchers considering research in a similar area, such as the adoption of new technologies in the Database as a Service (DaaS) industry. The insightful findings from this research can complement previously accumulated knowledge on MTD adoption and usage.

Methodological contributions made by this research include the factor that different data collection methods were used, to assist in increasing the validity of research findings. The research made use of triangulated data involving semi-structured interviews and the review of documents such as company reports as primary sources of evidence whilst secondary sources of evidence comprised mainly journals, conference papers, textbooks and so on.

This research again makes a substantial contribution from the research methodology, having established and validated measures relating to the different constructs of the research, including those in the framework. Methodologically, the research employed a questionnaire survey, qualitative interviews, observations and review of documents as data collection tools. This means that the research employed both qualitative and quantitative approaches in order to provide in-depth information about the subject.

9.5 Limitations of Research

Although the research has reached its aim and objectives, there were some unavoidable limitations which are itemized and discussed below.

One limitation of this research is the fact that the collection of empirical data depended mainly on the level of access that was granted to the researcher and the coverage of the questionnaires. Therefore, the participants could have hidden some vital information from the researcher, which could possibly have improved the research outcome, without the researcher's knowledge. This has limited the research findings to the data available to the researcher.

Another limitation during this study is the low response rate because of the time constraint. This research was conducted only with a very small sample size. This is due to the level of expertise of the participants expected to join in the survey. In the focus group with over 100 questionnaires administered, only 30 were returned, while in the second survey for the experts only 41 participants were recorded. Further research work would be undertaken to involve a large group of participants.

This study validation interview was limited to one user organisation of MTD in the United Kingdom, the residence of the researcher. It is the researcher's belief that although the validation interview was limited to UK, nevertheless, some of the findings are likely to be similar to those in other parts of the world. However, the present research validation results cannot be generalised without additional research. Similarly, despite the fact that issues concerning MTD all over the world are homogeneous, this is another opportunity for further research in this area.

9.6 Recommendations for further Research

The findings of this research and the research limitations have resulted in the identification of potential future research directions for investigation. The recommendations for further research as a result of this study are indicated below.

Further research could be done in the area of the framework, which should be validated in different contexts and other parts of the world to extend the generalisability and contribution of the framework. This can be achieved by re-testing the research findings and recommendations in other parts of the world, which will help to determine whether the findings are also well accepted as this.

Also, there could be further investigations that can extend the framework as new factors, such as the maintenance process which is not presently included in the framework, could emerge over time.

A team of researchers should be adopted in further studies where different persons will handle different locations. This will give room for comparative analysis of the different results in order to reach formidable and more generalised findings.

It appears from the literature review that there not much existing research that has examined the surrounding issues related to the adoption, utilization and maintenance of MTD concept. Thus, more research is still needed to be conducted in these areas to further expand the scope of the current research.

9.7 Summary

Future research is expected to extend the knowledge acquired in this research to other regions or areas not covered. The investigation in this research suggest that MTD is beneficial to organisations in that the technology amortises the total cost of ownership of DBMS, thereby leading to a reduction in per-tenant cost by increasing the scale of the operation. However,

there are salient factors that influence the adoption of MTDs. This research presents these factors with their individual impact on the decision of adoption. The acceptance level of MTDs is also one of the issues investigated in this research.

This research has fulfilled its goals and expectations by providing answers to all the research questions set out at the beginning of the study. Significant contributions have been provided towards explaining the impact of the factors influencing the evaluation and adoption of MTD.

This Chapter has presented the contributions of this research to the body of knowledge which includes the research methods adopted for the study and how they were applied; the developed framework; key limitations of the present research as well as recommendations for future research. The research also adds to the body of knowledge by empirically providing evidence that can increase the knowledge of MTD adoption, the influencing factors and the degree of impact in organisations thereby expanding the research area, in the field of IT. The research findings are beneficial to academics, intending users, users and service providers of DaaS.

REFERENCES

Abadi, D.J. (2012) Consistency tradeoffs in modern distributed database system design. Computer-IEEE Computer Magazine [online], **45**(2), pp. 37.

Abeysekera, S. (2001) Analysis approaches in participatory work involving ranks or scores. *DFID Theme Paper (revised).UK: Statistical Services Centre, University of Reading* [online].

Ahn, S. (2014) An evaluation of green manufacturing technologies based on research databases. *International Journal of Precision Engineering and Manufacturing-Green Technology* [online]**, 1**(1), pp. 5-9 .

Alkass, S., Mazerolle, M. and Harris, F. (1998) Rigour in research and peer-review: a reply. *Construction Management & Economics,* **16** (2), p. 139.

Al-Aqrabi, H., Liu, L., Hill, R. and Antonopoulos, N. (2015) Cloud BI: Future of business intelligence in the Cloud. *Journal of Computer and System Sciences* [online]**, 81**(1), pp. 85-96 Available at:<http://www.sciencedirect.com/science>.

Ambrose, K (2013) Use of Oracle 12c Pluggable Database: *Oracle Openworld 2013* [online]. [Accessed 23 October, 2013]. Available at <https://oracleus.activeevents.com/2013/connect/search.ww?eventRef=openworld>

Ankrah, N. A. (2007) An Investigation into the Impact of Culture on Construction Project Performance. PhD Thesis, University of Wolverhampton.

Aswathanarayana, H. (2011) Multi-tenant enabling a single-tenant application: *Wipro technologies* [online]. Bangalore: [Accessed 14 March 2013]. Available at <http://www.wipro.com>

Aulbach, S., Grust, T., Jacobs, D., Kemper, A. and Rittinger, J. (2008) Multi-tenant databases for software as a service: schema-mapping techniques *Proceedings of the 2008 ACM SIGMOD international conference on Management of data.* [online]. Vancouver, Canada New York, NY, USA: ACM, pp.1195-1206.

Aulbach, S., Jacobs, D., Kemper, A. and Seibold, M. (2009) A comparison of flexible schemas for software as a service *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data.* [online]. Providence, Rhode Island, USA New York, NY, USA: ACM, pp.881-888.

Aulbach, S., Seibold, M., Jacobs, D. and Kemper, A. (2011) Extensibility and Data Sharing in evolving multi-tenant databases *Data Engineering (ICDE), 2011 IEEE 27th International Conference on Data Engineering.* [online]. pp.99-110.

Avram, M.G. (2014) Advantages and challenges of adopting cloud computing from an enterprise perspective. Procedia Technology, 12, pp.529-534.

Bajaj, A. (1998) Factors relevant to senior information systems managers' decisions to adopt new computing paradigms: An exploratory study. Unpublished manuscript].Retrieved from http://www-afs.secure-endpoints.com/afs/andrew.cmu.edu/supa/wpapers/1998-1.pdf [online].

Bajaj, A. (2000) A study of senior information systems managers decision models in adopting new computing architectures. Journal of the AIS [online], 1(1es), pp. 4 .

Baker, M. J. (2003) Data Collection - Questionnaire Design. *The Marketing Review*, Vol. 3, pp. 343-370.

Banerjee, A., Ghosh, S. C. and Banerjee, N. (2012) Pack your sack for the cloud *Proceedings of the 5th India Software Engineering Conference.* [online]. Kanpur, India New York, NY, USA: ACM, pp.157-116.

Banville, R. and Holzel, R. (2012) Openedge multi-tenancy overview: *Progress software* [online]. Bedford: [Accessed 23 January 2013]. Available at < http://www.progress.com>

Bell, D.A. (2014) *Relational Databases: State of the Art Report 14: 5.* [online] Elsevier.

Bezemer,C. and Zaidman,A. (2010) Multi-tenant SaaS applications: maintenance dream or nightmare? *Proceedings of the Joint ERCIM Workshop on Software Evolution (EVOL) and International Workshop on Principles of Software Evolution (IWPSE)*. [online]. ACM, pp.88-92.

Bezemer, C., Zaidman, A., Platzbeecker, B., Hurkmans, T. and Hart, A. (2010) Enabling multi-tenancy: An industrial experience report *Software Maintenance (ICSM), 2010 IEEE International Conference on Software Maintenance.* [online]. pp.1-8.

Bouayad, A., Blilat, A., El Houda Mejhed, N. and El Ghazi, M. (2012) Cloud computing: Security challenges *Information Science and Technology (CIST), 2012 Colloquium in.* [online]. pp.26-31.

Boytsov, E. and Sokolov, V. (2012) The problem of creating multi-tenant database clusters. *Proceedings of the Spring/Summer Young Researchers' Colloquium on Software Engineering (SYRCoSE)* [online]**, 2012**.

Bleeker, S., Moll, H., Steyerberg, E., Donders, A., Derksen-Lubsen, G., Grobbee, D. and Moons, K. (2003) External validation is necessary in prediction research: A clinical example. *Journal of clinical epidemiology* [online], 56(9), pp. 826-832 .

Brinberg, D. and McGrath, J. E. (1985) *Validity and the Research Process*. Beverly Hills, CA: Sage Publications.

Bryman, A. (2008) *Social Research Methods*. 3rd ed. Oxford: Oxford University Press.

Burney S.M.A (2008) Inductive and Deductive Research Approach: Available online at http://www.drburney.net>

Burns, R.B (2000) Introduction to research methods (4th Edition). SAGE Publications, London.

Ceri, S. and Pelagatti, C.(1984) Distributed Databases: Principles and Systems. New York, NY: McGraw-Hill, 1984.

Ceri, S., Pernici, B. and Wiederhold, G. (1987) Distributed database design methodologies. Proceedings of the IEEE [online], **75**(5), pp. 533-546 .

Chong, R (2012) *Designing a Database* Castillo, E., Gutierrez, J. and Hadi, A. (2012) *Expert systems and probabilistic network models.* [online] Springer Science & Business Media.

For Multi-tenancy on the Cloud [online]. [Accessed 20 October, 2013]. Available at http://www.ibm.com/developerworks/data/library/techarticle>

Chong, R., Carraro, G., Wolter, R. (2006) *Multi-tenant Data Architecture* [online]. [Accessed 23 August 2013]. Available at :<http://www.msdn.microsoft.com/en-us/library/aa479086.aspx#mlttntda_topic3>

Corbett, J.C., Dean, J., Epstein, M., Fikes, A., Frost, C., Furman, J.J., Ghemawat, S., Gubarev, A., Heiser, C. and Hochschild, P. (2013) Spanner: Google's globally distributed database. ACM Transactions on Computer Systems (TOCS) [online], **31**(3), pp. 8 .

Coronel, C. and Morris, S. (2016) Database systems: design, implementation, & management. [online] Cengage Learning.

Creswell, J. (2003) Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd Ed), Sage, Thousand Oaks, CA.

Das, S., Nishimura, S., Agrawal, D. and El Abbadi, A. (2010) *Live database migration for elasticity in a multitenant database for cloud platforms* [online].

Davison, M. (2003) The legal protection of databases, Cambridge University Press, Cambridge.

Denscombe, M. (2007) *The Good Research Guide for small-scale social research projects.* Berkshire: Open University Press.

De Vos, A. S., Strydom, H., Fouche, C. B. and Delport, C. S. L. (2002) *Research at grass roots for the social sciences and human professions*. Pretoria: Van Schaik.

Dodge, Y. (2006) The Oxford dictionary of statistical terms. [online] Oxford University Press.

Easterby-Smith, M., Thorpe, R. and Lowe, A. (1991) *Management Research: An Introduction*. London. Sage Publications Ltd.

Easterby-Smith M, Thorpe, R and Lowe, A (2002): Management Research: an introduction; 2nd edition; Sage.

Egbu, A. U. (2007) Impact of Land Use Planning on Urban Housing Development in Nigeria. PhD Thesis, University of Wolverhampton.

Eisenhardt, K. M. (1989) Building theories from case study research. Academy of Management Review, 14(4), pp.532-550.

Eisenhardt, K. M. and Howe K. R. (1992) Validity in educational research. In LeCompte, M.D. Millroy, W. L. and Preissle, J. (Eds.). The handbook of qualitative research in education (pp. 643-680). San Diego, CA: Academic Press.

El Abbadi, A and Toueg, S. (1986) "Availability in partitioned replicated databases," in Proc. 5th ACM SICACT-SlCMOD Symp. on Principles of Database Systems, pp. 240-251, 1986.

Elmore, A. J., Das, S., Agrawal, D. and Abbadi, A. (2011) Towards an elastic and autonomic multitenant database *Proc. of NetDB Workshop.* [online].

Elmore, A. J., Das, S., Agrawal, D. and El Abbadi, A. (2011) Zephyr: Live Migration in Shared Nothing Databases for Elastic Cloud Platforms *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data.* [online]. Athens, Greece New York, NY, USA: ACM, pp.301-312.

Ezejejue, C and Ogwu, E (1990): Basic Principles in Managing Research Project, African-Feb Publishers limited

Fabling, R. and Sanderson, L. (2016) A Rough Guide to New Zealand's Longitudinal Business Database. *Available at SSRN* [online].

Fang, S. and Tong, Q. (2011) A comparison of multi-tenant data storage solutions for Software-as-a-Service *Computer Science and Education (ICCSE), 2011 6th International Conference on.* [online]. pp.95-98.

Fellows, R. and Lui, A. (1997) *Research Methods for Construction*. Blackwell Science, Oxford: Maiden, MA, USA.

Fellows, R., Langford, D. Newcombe, R. and Urry, S. (2002) Construction Management in Practice. 2nd ed. Oxford. Blackwell Science Ltd, UK.

Fellows, R. and Liu, A. (2003) Research Methods for Construction. Blackwell Ltd, UK.

Fenn, P. (1997) Rigour in research and peer review, *Construction Management & Economics,* **15** (4), p. 383.

Flood, J. and Keane, A. (2012) A Proposed Framework for the Active Detection of Security Vulnerabilities in Multi-tenancy Cloud Systems *Emerging Intelligent Data and Web Technologies (EIDWT), 2012 Third International Conference on Emerging Intelligent Data and Web Technologies.* [online]. pp.231-235.

Frees, W. E. (1996) *Data Analysis Using Regression Models: The Business Perspective*. Prince Hall, Englewood Cliffs, USA.

Fuchs, M., Bossert, D. and Stukowski, S. (2013) Response rate and nonresponse bias-impact of the number of contact attempts on data quality in the European Social Survey. *Bulletin of*

Sociological Methodology/Bulletin de Méthodologie Sociologique [online]**, 117**(1), pp. 26-45 .

Gall, D. M., Walker, R. B. and Joyce, P. G. (1996) Educational Research: an introduction. 6th Ed. White Plains, N.Y.: Longman.

Gangwar, H., Date, H. and Ramaswamy, R. (2015) Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *Journal of Enterprise Information Management* [online]**, 28**(1), pp. 107-130 .

Gao, B., An, W., Sun, X., Wang, W., Fan, L., Guo, C et al. (2011) A Non-intrusive Multi-tenant Database Software for Large Scale SaaS Application *e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on e-Business Engineering.* [online]. pp.324-328.

Garcia-Molina, H. and Abbott, R.K. (1987) Reliable distributed database management. Proceedings of the IEEE [online], **75**(5), pp. 601-620.

Gessert,F., Bücklers,F. and Ritter,N. (2014) Orestes: A scalable Database-as-a-Service architecture for low latency *Data Engineering Workshops (ICDEW), 2014 IEEE 30th International Conference on Data Engineering Workshops.* [online]. IEEE, pp.215-222.

Gey, F., Van Landuyt, D., Joosen, W. and Jonckers, V. (2015) Continuous evolution of multi-tenant SaaS applications: A customizable dynamic adaptation approach. *status: accepted* [online].

Giarratano, J. and Riley, G. (1998) *Expert systems: principles and programming.* [online] 3rd ed. Boston, Mass: PWS.

Golafshani, N. (2003) Understanding Reliability and Validity in Qualitative Research. The Qualitative Report, **8**(4), pp. 597-607.

Gordon, S. and Gordon, J. (1991) Organisational determinants of distributed database management systems (DDBMS) adoption *Proceedings of the 1991 Information Resources Management Association international conference on Managing information technology in a global society*. [online]. IGI Global, pp.267.

Gordon, S. and Gordon, J. (1993) Factors that affect the adoption of distributed database management systems *Proceedings of the 1993 conference on Computer personnel research*. [online]. ACM, pp.151-167.

Grosheide, F.W. (2002) Database Protection--The European Way. *Wash.UJL and Pol'y* [online]**, 8**pp. 39 .

Grund, M., Schapranow, M., Krueger, J., Schaffner, J. and Bog, A. (2008) Shared Table Access Pattern Analysis for Multi-Tenant Applications *Advanced Management of Information for Globalized Enterprises, 2008. AMIGE 2008. IEEE Symposium on Advanced Management of Information for Globalized Enterprises.* [online]. pp.1-5.

Gündüz, M., Nielsen, Y. and Özdemir, M. (2013) Quantification of delay factors using the relative importance index method for construction projects in Turkey. *Journal of Management in Engineering* [online]**, 29**(2), pp. 133-139 .

Gunduz, M., Nielsen, Y. and Ozdemir, M. (2015) Fuzzy assessment model to estimate the probability of delay in Turkish construction projects. *Journal of Management in Engineering* [online]**, 31**(4), pp. 04014055 .

Guo, C., Sun, W., Huang, Y., Wang, Z. and Bo Gao. (2007) A Framework for Native Multi-Tenancy Application Development and Management *E-Commerce Technology and the 4th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, 2007. CEC/EEE 2007. The 9th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services.* [online]. pp.551-558.

Hacigümüs,H., Iyer,B. and Mehrotra,S. (2002) Providing database as a service *Data Engineering, 2002. Proceedings. 18th International Conference on Data Engineering.* [online]. IEEE, pp.29-38.

Hair, J. F., Anderson, R. E., Tatham, R. L. and Black, W. C. (1998) *Multivariate data analysis*. Upper Saddle River, N.J., Prentice Hall.

Hanna, E. M., Mohamed, N. and Al-Jaroodi, J. (2012) The Cloud: Requirements for a Better Service *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012).* [online]. Washington, DC, USA: IEEE Computer Society, pp.787-792.

Hinton, P.R., McMurray, I. and Brownlow, C. (2014) *SPSS explained.* [online] Routledge.

Holt, G. (2013) Asking questions, analysing answers: relative importance revisited. *Construction Innovation* [online]**, 14**(1), pp. 2-16 .

Hui, M., Jiang, D., Li, G. and Zhou, Y. (2009) Supporting Database Applications as a Service *Data Engineering, 2009. ICDE '09. IEEE 25th International Conference on Data Engineering.* [online]. pp.832-843.

Jabareen, Y.R. (2009) Building a conceptual framework: philosophy, definitions, and procedure. *International Journal of Qualitative Methods* [online]**, 8**(4), pp. 49-62 .

Jackson, P. (1999) *Introduction to expert systems.*[online] Harlow: Addison-Wesley. Available at: <http://wlv.summon.serialssolutions.com>

Jacobs, D. and Aulbach, S. (2007) Ruminations on multi-tenant databases. *BTW Proceedings* [online]**, 103**pp. 514-521 .

Jasti, A., Shah, P., Nagaraj, R. and Pendse, R. (2010) Security in multi-tenancy cloud *Security Technology (ICCST), 2010 IEEE International Carnahan Conference on Security Technology.* [online]. pp.35-41.

Johnson, C. and Grandison, T. (2007) *Compliance with data protection laws using Hippocratic Database active enforcement and auditing* [online].

Kaur, K and Singh, H (2016) Distributed database system on web server: A Review. *International Journal of Computer Techniques* [online], **3**(6)

Keemti, P. (2010) *Multi-tenant Database Architecture* [online]. [Accessed 23 August 2013]. Available at <http://www.msdn.microsoft.com/eus/library/aa479086.aspx#mlttntda_topic1>

Kennedy R. C., Xiang, X., Madey, G. R., and Cosimano, T. F. (2005) Verification and validation of scientific and economic models. Agent 2005 Conference Proceedings. Chicago, IL, October 2005.

Khan, M.F. and Ullah, M.A. (2012) An Approach Towards Customized Multi-Tenancy. *International Journal of Modern Education and Computer Science (IJMECS)* [online]**, 4**(9), pp. 39 .

Kim, W. (2009) Cloud Computing: Today and Tomorrow. Journal of object technology [online], **8**(1), pp. 65-72 .

Kim,W., Kim,S. D., Lee,E. and Lee,S. (2009) Adoption issues for cloud computing *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*. [online]. ACM, pp.2-5.

Kirk, J., and Miller, M. L. (1986) *Reliability and validity in qualitative research.* Beverly Hills: Sage Publications.

Khajeh-Hosseini, A., Greenwood, D., Smith, J.W. and Sommerville, I. (2012) The cloud adoption toolkit: supporting cloud adoption decisions in the enterprise. Software: Practice and Experience [online], **42**(4), pp. 447-465.

Kuhlenkamp, J., Klems, M. and Röss, O. (2014) Benchmarking scalability and elasticity of distributed database systems. *Proceedings of the VLDB Endowment*, **7**(12), pp.1219-1230.

Kshetri, N. (2013) Privacy and security issues in cloud computing: The role of institutions and institutional evolution. Telecommunications Policy [online], **37**(4), pp. 372-386.

Kumar, N., Bilgaiyan, S. and Sagnika, S. (2013) An Overview of Transparency in Homogeneous Distributed Database System. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* Volume [online], 2.

Lang, W., Shankar, S., Patel, J. M. and Kalhan, A. (2012) Towards Multi-tenant Performance SLOs *Data Engineering (ICDE), 2012 IEEE 28th International Conference on.* [online]. pp.702-713.

Lucko, G. and Rojas, E.M. (2010) Research Validation: Challenges and Opportunities in the Construction Domain. *Journal of Construction Engineering and Management* [online]**,** **136**(1), pp. 127-135 .

Luo, Y., Zhou, S. and Guan, J. (2015) LAYER: A cost-efficient mechanism to support multi-tenant database as a service in cloud. *Journal of Systems and Software* [online]**, 101**(0), pp. 86-96 Available at:<http://www.sciencedirect.com/science>.

Maher,M. (1987) Expert systems for civil engineers: technology and application [online]. ASCE, .

Matthew, O., Buckley, K., Garvey, M and Moreton, R (2016c) Multi-Tenant Database Framework Validation and Implementation into an Expert System. *International Journal of Advanced Studies in Computer Science and Engineering* **5**(8)

Matthew, O., Buckley, K., and Garvey, M (2016b) Statistical Analysis of Factors That Influences the Evaluation and Adoption of Multi-Tenant Databases *International. Journal of Computer Trends and Technology* **37**(2), pp.85-95

Matthew, O., Buckley, K., and Garvey, M (2016a) A Framework for Multi-Tenant Database Adoption Based on the Influencing Factors. *International Journal of Information Technology and Computer Science* **8**(3), pp.1-9.

Matthew, O., Buckley, K., and Garvey, M (2015) Predicting the Impact of the Factors That Influence The Adoption Of Multi-Tenant Databases. *International Conference on Computer and Information Science and Technology (CIST'15)*, University of Ottawa, Ottawa, Canada; 05/2015

Matthew, O., Dudley, C and Moreton, R (2014) A Review of Multi-Tenant Database and Factors That Influence Its Adoption. *UKAIS 2014 Conference St* Catherine's College, University of Oxford, 7-9th April. Oxford, Oxford, UK; 04/2014

Maxwell, J. A. (1992) Understanding and validity in qualitative research. *Harvard Educational Review*, **62**, pp. 279-300.

McDermott, R. (2011) Internal and external validity. *Cambridge handbook of experimental political science* [online]pp. 27-40 .

Mcgoo.com (2013) Building an Expert System in ES-Builder [online] Available at:www.mcgoo.com.au/esbuilder/index.php

McGregor, S. L., and Murname, J. A. (2010). Paradigm, methodology and method: Intellectual integrity in consumer scholarship. *International Journal of Consumer Studies*, ***34*** (4), 419-427.

McKendrick, J. (2016) Database as a service enters the enterprise mainstream [online] Available at:http://www.unisphereresearch.com [Accessed January, 2017]

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded source book
(2nd ed.). Newbury Park, CA: Sage.

Miller, R. and Acton, C. (2009) *SPSS for social scientists.* [online] Palgrave Macmillan.

Minkarah, I. and Ahmad, I. (1989) Expert systems as construction management tools. *Journal of Management in Engineering* [online]**, 5**(2), pp. 155-163 .

Mintzberg, H. (1973) The Nature of Managerial Work, Harper and Row, New York.

Morse J. M., Hupcey, J. E., Penrod, J., Spiers, J. A., Pooler, C., & Mitcham, C. (2002). Symposium conclusion: Issues of validity—Behavioral concepts, their derivation and interpretation. International Journal of Qualitative Methods, **1**(4), 68–73.

MTD Validation Survey (2016) viewed 31 May 2016 [online] Available at <https://docs.google.com/forms/d/14eke2QAUJog8HhIgssH_TMli9CUeNKgrQ2KivuIgKCY/edit

Narasayya,V. R., Das,S., Syamala,M., Chandramouli,B. and Chaudhuri,S. (2013) SQLVM: Performance Isolation in Multi-Tenant Relational Database-as-a-Service.*CIDR.* [online].

Ni, J., Li, G., Zhang, J., Li, L. and Feng, J. (2012) Adapt: adaptive database schema design for multi-tenant applications *Proceedings of the 21st ACM international conference on Information and knowledge management.* [online]. Maui, Hawaii, USA New York, NY, USA: ACM, pp.2199-2203.

Ni, J., Li, G., Wang, L., Feng, J., Zhang, J. and Li, L. (2014) Adaptive Database Schema Design for Multi-Tenant Data Management. [online].

O'Brien, J. and Marakas, G.M.(2008) Management Information Systems (pp. 185-189). New York, NY: McGraw-Hill Irwin

Oppenheim, A. N. (1992). Questionnaire Design, Interviewing and Attitude Measurement, Pinter, London.

Özsu, M.T. and Valduriez, P. (2011) Principles of distributed database systems. Springer Science & Business Media.

Pandithurai, O., Poongodi, M., Kumar, S. P. and Krishnan, C. G. (2011) A method to support multi-tenant as a service *Advanced Computing (ICoAC), 2011 Third International Conference on Advanced Computing.* [online]. pp.157-162.

Pengcheng Xiong, Yun Chi, Shenghuo Zhu, Hyun Jin Moon, Pu, C. and Hacgumus, H. (2015) SmartSLA: Cost-Sensitive Management of Virtualized Resources for CPU-Bound Database Services. *Parallel and Distributed Systems, IEEE Transactions on Parallel and Distributed Systems* [online]**, 26**(5), pp. 1441-1451 .

Pippal, S.K. and Kushwaha, D.S. (2013) A simple, adaptable and efficient heterogeneous multi-tenant database architecture for ad hoc cloud. *Journal of Cloud Computing* [online]**, 2**(1), pp. 1-14 .

Pippal, S., Sharma, V., Mishra, S. and Kushwaha, D. (2011) Secure and efficient multitenant database for an ad hoc cloud *Securing Services on the Cloud (IWSSC), 2011 1st International Workshop on Securing Services on the Cloud.* [online]. pp.46-50.

Rahimi, S.K. and Haug, F.S. (2010) Distributed database management systems: A Practical Approach. [online] John Wiley & Sons.

Rea L. M. and Parker, P. A. (1997). *Designing and Conducting Survey Research*, 2nd Ed., Jossey-Bass Publishers, San Francisco, USA.

Rea, L.M. and Parker, R.A. (2014) *Designing and conducting survey research: A comprehensive guide.* [online] John Wiley & Sons.

Rosenthal, R. and Rosnow, R. L. (1991) Essentials of Behavioral Research: Methods and Data Analysis. New York; London, McGraw-Hill.

Runeson, G. R. and Loosemore, M. (1999) Gate-keepers or Judges: Peer-Reviews in Construction Management, *Journal of Construction Management and Economics*, **17**(4), p. 529.

Sambasivan, M. and Soon, Y.W. (2007) Causes and effects of delays in Malaysian construction industry. *International Journal of Project Management* [online]**, 25**(5), pp. 517-526 Available at:<http://www.sciencedirect.com/science/article/pii/S0263786306001700>.

Sang, C., Li, Q. and Kong, L. (2012) Tenant Oriented Lock Concurrency Control in the Shared Storage Multi-tenant Database *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2012 IEEE 16th International Enterprise Distributed Object Computing Conference Workshops.* [online]. pp.179-189.

Sauermann, H. and Roach, M. (2013) Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features. *Research Policy* [online]**, 42**(1), pp. 273-286 .

Saunders M, Lewis P and Thornhill A (2007): Research methods for business students. 4th Edition, Prentice Hall

Saunders M, Lewis P and Thornhill A (2009): Research methods for business students. 5th Edition, Pearson Education Limited

Schaffner, J. (2013), Multi tenancy for cloud-based in-memory column databases: workload management and data placement, Springer, Wien

Schaffner, J., Jacobs, D., Kraska, T. and Plattner, H. (2012) The Multi-Tenant Data Placement Problem *DBKDA 2012, The Fourth International Conference on Advances in Databases, Knowledge, and Data Applications.* [online]. pp.157-162.

Schellekens, M. (2011) A database right in search results? – An intellectual property right reconsidered in respect of computer generated databases. *Computer Law and Security Review* [online]**, 27**(6), pp. 620-629 Available at:<http://www.sciencedirect.com>

Schiller, O., Schiller, B., Brodt, A. and Mitschang, B. (2011) Native support of multi-tenancy in RDBMS for software as a service *Proceedings of the 14th International Conference on Extending Database Technology.* [online]. Uppsala, Sweden New York, NY, USA: ACM, pp.117-128.

Shue, D., Freedman, M.J. and Shaikh, A. (2013) Fairness and isolation in multi-tenant storage as optimization decomposition. *SIGOPS Oper.Syst.Rev.* [online]**, 47**(1), pp. 16-21 .

Silverman, D. (2006) *Interpreting Qualitative Data.* 3rd ed. London: Sage.

Simons, H. (2009) Case study research in practice. Sage, London.

Singh, Y.K. and Bajpai, A.B. (2008) “Research Methodology: Techniques and Trends” APH Publishing Corporation Students” 4th edition, Prentice Hall

Sotto, L.J., Treacy, B.C. and McLellan, M.L. (2010) Privacy and Data Security Risks in Cloud Computing. *World Communications Regulation Report* [online]**, 5**(2), pp. 38 .

Sousa, F. and Machado, J. (2012) Towards Elastic Multi-Tenant Database Replication with Quality of Service *Utility and Cloud Computing (UCC), 2012 IEEE Fifth International Conference on Utility and Cloud Computing.* [online]. pp.168-175.

Stake, R. (1995). The art of case research. Sage Publications: California.

Sun, X., Zimmermann, C., Jackson, G., Bunker, C. and Harrington, P (2011) Classification of jet fuels by fuzzy rule-building expert systems applied to three-way data by fast gas chromatography—fast scanning quadrupole ion trap mass spectrometry. *Talanta* [online]**, 83**(4), pp. 1260-1268 Available at:<http://www.sciencedirect.com>.

Sutter, H.P. (2002) *Independent distributed database system* [online].

Sweet, S. A. and Grace-Martin, K. (2003) *Data Analysis with SPSS: A First Course in Applied Statistics.* 3rd ed. Allyn and Bacon, Boston.

Taylor, G. R. (2005) *Integrating quantitative and qualitative methods in research*. 2nd ed. University Press of America.

Tomar, P. and Megha (2014) An overview of distributed databases. *Intenational Journal of Information and Computation Technology* [online], **4**(2), pp.207-214.

Tomar, P. and Suruchi (2016) Efficient concurrency control mechanism for distributed databases *Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on Computing for Sustainable Global Development.* [online]. IEEE, pp.3415-3418.

Tonidandel, S. and LeBreton, J. (2011) Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology* [online]**, 26**(1), pp. 1-9 .

USA: Oracle multitenant for Dummies. (2013). Hoboken: John Wiley and Sons, Inc.

Vashistha, A and Ahmed, P. (2012)"SaaS Multi-Tenancy Isolation Testing- Challenges and Issues", *International Journal of Soft Computing and Engineering,*[online] **2**(5), pp. 49-50 pp. 49-50 Available at:<http://wlv.summon.serialssolutions.com>

Walraven, S., Van Landuyt, D., Truyen, E., Handekyn, K. and Joosen, W. (2014) Efficient customization of multi-tenant Software-as-a-Service applications with service lines. *Journal of Systems and Software* [online] (0), Available at :< http://www.sciencedirect.com/science>.

Wang, Z., Guo, C., Gao, B., Sun, W., Zhang, Z and An, W. (2008) A Study and Performance Evaluation of the Multi-Tenant Data Tier Design Patterns for Service Oriented Computing *e-Business Engineering, 2008. ICEBE '08. IEEE International Conference on e-Business Engineering.* [online]. pp.94-101.

Weis, J. and Alves-Foss, J. (2011) Securing database as a service: Issues and compromises. *IEEE Security & Privacy* [online](6), pp. 49-55 .

Wheeler, P (2013) Simplify Consolidation with Oracle Database 12c: *Oracle Openworld 2013* [online]. [Accessed 23 October, 2013]. Available at <https://oracleus.activeevents.com/2013/connect/search.ww?eventRef=openworld>

Wheeler, P and Michalewicz, M (2013) Hardware and Software Engineered To work Together: *Oracle Openworld 2013* [online]. [Accessed 23 October, 2013]. Available at <http: https://oracleus.activeevents.com/2013/connect/search.ww?eventRef=openworld>

Wijesudera, A. and Harris, F. (1986) Using expert systems in construction. *Construction Computing, July* [online].

Wood, K. and Anderson, M. (2011) Understanding the Complexity Surrounding Multitenancy in Cloud Computing *e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on e-Business Engineering.* [online]. pp.119-124.

Xiao, H. (2002) A Comparative Study of Contractor Performance based on Japanesse, UK and US Construction Practice. PhD Thesis, University of Wolverhampton

Yaish,H. and Goyal,M. (2013) A multi-tenant database architecture design for software applications *Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on Computational Science and Engineering.* [online]. IEEE, pp.933-940.

Yaish, H., Goyal, M. and Feuerlicht, G. (2011) An Elastic Multi-tenant Database Schema for Software as a Service *Dependable, Autonomic and Secure Computing (DASC), 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing.* [online]. pp.737-743.

Yaish, H., Goyal, M. and Feuerlicht, G. (2013) Proxy service for multi-tenant database access. *in Availability, Reliability, and Security in Information Systems and HCI.* Springer, .

Yang Z, Wang X and Su C (2006): A review of research methodologies in international business; international business review, vol. 15, pp. 601-617 available online at http://www.sciencedirect.com >

Ye, N. and Wu, T. (2014) *Developing Windows-based and Web-enabled Information Systems.* [online] CRC Press.

Yeboah-Boateng, E.O. and Essandoh, K.A., 2014. Factors influencing the adoption of cloud computing by small and medium enterprises in developing economies. *International Journal of Emerging Science and Engineering*, ***2***(4), pp.13-20.

Yin, R. K. (2002) Case Study Research, Design and Methods. 3rd ed. Sage Publications: London.

Yin, R.K. (2013) Case study research: Design and methods. [online] Sage publications.

Ying, H., Wang, Q., Wang, Z., and Wang, N. (2011) DB2MMT: A Massive Multi-tenant Database Platform for Cloud Computing *e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on e-Business Engineering.* [online]. pp.335-340.

Yu, H. and Wang, D. (2011) A heuristic data allocation method for multi-tenant SaaS application in distributed database systems *2011 International Conference on Information Management, Innovation Management and Industrial Engineering*. [online]. IEEE, pp.382-386.

Zhou, Y., Wang, Q., Wang, Z and Wang, N. (2011) DB2MMT: A Massive Multi-tenant Database Platform for Cloud Computing *e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on e-Business Engineering.* [online]. pp.335-340.

451 Research (2016) Database as a Sevice (DBaaS):Use cases and adoption patterns [online] Available at <http://www.oracle.com/us/products/database/451-report-dbaas-3038788> [Accessed January 2017]

APPENDICES

Appendix A – Cover Letter and Questionnaire for 1st Survey

Dear Participant,

SCIENTIFIC GUIDELINE TOWARDS MULTI-TENANT DATABASE DRIVE

My name is Olumuyiwa Matthew, a PhD student of the University of Wolverhampton, United Kingdom. As part of my programme, I am carrying out a survey and your participation will be appreciated. Therefore, I would like to invite you to help my research by completing a simple questionnaire.

Completion of the attached questionnaire will take about 10-15 minutes, and all questions can be answered by following the instructions. Completion of the questionnaire is completely voluntary. All responses are anonymous, there are no correct or incorrect answers, and respondents who take part will not be identifiable. If the results of this study are published they will only be a summary of all responses to ensure that privacy is protected. Returning this questionnaire will be considered as consent to participate in the survey.

A summary of findings will be available at the conclusion of the study and if you wish to obtain a copy of the results, please provide your contact details. Please note that all data obtained for this research will be stored securely for future references.

Thank you. Your participation and contribution are greatly appreciated.

Yours sincerely,

OLUMUYIWA . MATTHEW

PHONE: [mobile numbers redacted]

EMAIL: [e-mail addresses redacted]

CONCEPT OF MULTI-TENANCY IN DATABASE

A multi tenancy database refers to a principle where single instance of the DBMS runs on a server, serving multiple clients organisation (tenants). Multi-tenant database is one which provides database support to a number of separate and distinct groups of users, also referred to as tenants. A tenant is simply any logically defined group of users that requires access to its own set of data. This is an architectural pattern in which a single instance of the software is run on the service provider's infrastructure, and multiple tenants access the same instance. This concept provides the ability of a system to provide database management services to different users or customers without having interference with each other's processes. This reduces effort made in production and the cost incurred in the development.

In a multi-tenant enabled service environment, user requests from different organisations and companies (*tenants*) are served concurrently by one or more hosted application instances and databases based on a scalable, shared hardware and software infrastructure. Such database system must be able to maintain or even increase its performance or efficiency level under larger operational demands.

And multi-tenant data management is a major application of SaaS. Today many companies want to outsource their data to a third party which hosts a multi-tenant database system to provide data management service. Each company is called a tenant. The multi-tenant data management system amortizes the cost of hardware, software and professional services to a large number of tenants and thus significantly reduces per- tenant cost by increasing the scale. Thus the multi-tenant database system requires having excellent performance, low-space requirement and good scalability.

Multi-tenancy database system is a new technology that can be implemented in both host based and cloud based environment. My research will be focusing on the trend in this technology, examining the security policies and loopholes in the implementation of the technology with a view to producing a standard scientific guideline for the drive towards database multi-tenancy.

Section I: Personal Information

Kindly tick the correct option in the box provided.

1. In what age group are you?

- ○ 19 and under
- ○ 20 - 29
- ○ 30 - 39
- ○ 40 - 49
- ○ 50 - 59
- ○ 60 +

2. Gender:

- ○ Male
- ○ Female

3. Name of Organisation…….

…….

4. Address of Organisation

……
……

5. In terms of your current occupation, how would you characterize yourself?

- ☐ Writer
- ☐ Administrative Assistant
- ☐ Journalist
- ☐ Secretary
- ☐ Academic
- ☐ Professional
- ☐ Technical expert

☐ Student

☐ Designer

☐ Administrator/Manager

☐ Other, please specify below:

Section 2

Kindly tick your answer in the box provided. All factors are given equal weight from question 11

6. Do you work with any database system?

Yes ☐

No ☐

7. How do you classify your level of involvement in database technology?

Novice ☐ Less than 6months.

Intermediate ☐ Between 6 months - 3years.

Expert ☐ More than 3 years.

8. If more than 3years, are you an Administrator?

Yes ☐ No ☐

9. Are you currently in use of a multi-tenants database?

Yes ☐ No ☐

10. How many tenants are on this multi-tenant database?

Below 10 ☐ 11-20 ☐

21-30 ☐ 31- 40 ☐

41-50 ☐ 50 above ☐

11. Do you believe that the following factors contribute to the drive towards multi-tenant database? Tick as many options that is applicable.

Cost ☐ Security ☐

Size of tenant database ☐ Data Isolation ☐

Number of tenants ☐ Number of users per tenant ☐

Growth rate of number of tenants ☐ Growth rate of tenant database ☐

Regulations ☐

12. What are the other factors which you are aware that are not mentioned above that also contribute to drive towards multi-tenancy? Kindly list them below.

13. What do you think is the degree of influence that each of these factors has on the drive towards multi-tenancy database? Ranking from 1-5 (5- very influential; 4- influential; 3- mild; 2- very mild; 1- no influence)

Cost ☐ Security ☐

Size of tenant database ☐ Data Isolation ☐

Number of tenants ☐ Number of users per tenant ☐

Growth rate of number of tenants ☐ Growth rate of tenant database ☐

Regulations ☐ Others – (mention in 11) ☐

14. In your personal opinion, In the last 2 years, how has the acceptability or acceptances of multi-tenant database changed?

Highly Improved. ☐ Improved. ☐

No difference ☐ Reduced. ☐

Highly Reduced . ☐

15. Do you plan to use oracle Multi-tenant database option?

Yes ☐ No ☐

If Yes go to 16 and if No go to 17.

16. Is it with one pluggable database (PDB) or more?

One PDB ☐ More than one PDB ☐

17. Which other option(s) of Multi-tenant database are you using or like to use? List them below.

18. From your options in 17 above, how can you rate their level of usage? Ranking from 1-5 (5- very high; 4- high; 3- low; 2- very low; 1- poor)

Option 1 ☐ Option 2 ☐ Option 3 ☐

Option 4 ☐ Option 5 ☐

THANK YOU FOR YOUR PARTICIPATION.

Appendix B - Cover Letter for 2nd Survey

Dear Participant,

SCIENTIFIC GUIDELINE TOWARDS MULTI-TENANT DATABASE DRIVE

My name is Olumuyiwa Matthew, a PhD student of the University of Wolverhampton, United Kingdom. As part of my programme, I am carrying out a survey and your participation will be appreciated. Therefore, I would like to invite you to help my research by completing a simple questionnaire.

Completion of the survey will take about 10-15 minutes, and all questions can be answered by following the instructions. Completion of the questionnaire is completely voluntary. All responses are anonymous, there are no correct or incorrect answers as all responses are based on your opinion about the subject area, and respondents who take part will not be identifiable. If the results of this study are published they will only be a summary of all responses to ensure that privacy is protected.

A multi-tenant database refers to a principle where a single instance of the DBMS runs on a server, serving multiple client organisations (tenants). A Multi-tenant database is one which provides database support to a number of separate and distinct groups of users, also referred to as tenants. A tenant is simply any logically defined group of users that requires access to its own set of data. This is an architectural pattern in which a single instance of the software is run on the service provider's infrastructure, and multiple tenants access the same instance. This concept provides the ability of a system to provide database management services to different users or customers without having interference with each other's processes. This reduces effort made in production and the cost incurred in the development. And multi-tenant data management is a major application of SaaS.

My research will be focusing on the trends in this technology, examining the security policies and loopholes in the implementation of the technology with a view to producing a standard scientific guideline for the drive towards multi-tenant databases. Here is the link to the survey, copy and paste on a browser - http://survey.wlv.ac.uk/survey.asp?s=01044001164086125187

A summary of findings will be available at the conclusion of the study and if you wish to obtain a copy of the results, please contact me through my details below. Please note that all data obtained for this research will be stored securely for future references.

Thank you. Your participation and contribution are greatly appreciated.

Yours sincerely,

OLUMUYIWA . MATTHEW

PHONE: [mobile numbers redacted]

EMAIL: [e-mail addresses redacted]

Appendix C - Questionnaire for 2nd Survey

Survey input field Respondent's answer

Name:

Profession:

Age:

Email Address:

Country:

Name of Organisation:

Type of Organisation:

1. Do you work with any database system or have you ever worked with one before?

1. Yes
2. No

2. How do you classify your level of involvement in database technology?

1. Novice (Less than 6 months)
2. Intermediate (Between 6 months - 3 years)
3. Expert (More than 3 years)

3. Are you a database administrator?

1. Yes
2. No

4. Are you aware of the concept called "MULTI-TENANT DATABASE" (MTD)?

1. Yes
2. No

5. Are you currently in use of Multi-tenant database (MTD) or have you used one before?

1. Yes
2. No

6. How many tenants are/ were on this multi-tenant database (MTD) environment?

1. Below 10
2. 11-20
3. 21-30
4. 31-40
5. 41-50
6. 50 above

7. Are you a MTD provider or user?

1. Provider
2. User

8. "FACTORS AFFECTING THE ADOPTION OF A MTD"

Time is said to be the total length of period in second/minutes/ hours/days it will take to build and configure MTD.
Do you believe that time is one of the factors affecting the adoption of a MTD?

1. Yes
2. No

9. What do you think is the degree of effect time has on the adoption of MTD?
Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

10. Cost is the total cost of ownership which is broken down into infrastructural cost, management cost and application development cost. Do you believe that cost is one of the factors affecting the adoption of a MTD?

1. Yes
2. No

11. What do you think is the ranking of these three categories of Cost in the process of building a MTD?

	LOW	MEDIUM	HIGH
INFRASTRUCTURAL COST	-	-	-
MANAGEMENT COST	-	-	-
APLLICATION DEVELOPMENT COST	-	-	-

12. What is the degree of effect cost has on the adoption of MTD?
Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

13. Data Isolation is the ability of MTD to ensure that data of one tenant is isolated from another tenants.
Do you believe that data isolation is one of the factors affecting the adoption of MTD?

1. Yes
2. No

14. What is the degree of effect data isolation has on the adoption of a MTD?
Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4

5. Very Influential - 5

15. Scalability means that database must have the capability of on-demand scale to support large volumes of tenants. Do you believe that scalability is one of the factors affecting the adoption of MTD?

1. Yes
2. No

16. What do you think is the degree of effect scalability has on the adoption of MTD? Ranking from 1 - 5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

17. Flexibility in MTD is the ability to extend the base schema to support multiple tenants and to serve thousands of tenants through one instance.
Do you believe that flexibility is one of the factors affecting the adoption of MTD?

1. Yes
2. No

18. What is the degree of effect flexibility has on the adoption of MTD?
Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

19. Customization is changing standard application functionalities to suit individual business requirement without imparting on other tenants.
Do you believe that customization is one of the factors affecting the adoption of a MTD?

1. Yes
2. No

20. What is the degree of effect customization has on the adoption of a MTD?
Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

21. Regulations are laws and policies put in place by different governments that serve as protection to databases of different entities.
Do you believe that regulations is one of the factors affecting the adoption of a MTD?

1. Yes

2. No

22. What is the degree of effect regulations has on the adoption of MTD?
Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

23. The size of a tenant's database is the capacity of each tenant database. Do you believe the size of tenant's database is one of the factors affecting the adoption of MTD?

1. Yes
2. No

24. What is the degree of effect the size of tenant's database has on the adoption of MTD? Ranking from 1-5.

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

25. The number of tenants is the total number of independent tenants on a particular MTD. Do you believe that the number of tenants is one of the factors affecting the adoption of MTD?

1. Yes
2. No

26. How can you now rank the degree of influence the number of tenants has on the adoption of MTD? Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

27. The number of users per tenant is the number of active people under each tenant's account that use the database. Do you believe that this could be one of the factors that affect the adoption of MTD?

1. Yes
2. No

28. What do you think is the degree of effect the number of users per tenant has on the adoption of MTD? Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3

4. Influential -4
5. Very Influential - 5

29. The growth rate of number of tenants describes the rate at which the number of independent tenants increases on the MTD. Do you believe that this could be one of the factors affecting the adoption of MTD?

1. Yes
2. No

30. What do you think will be the degree of effect the growth rate of number of tenants has on the adoption of MTD? Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

31. The growth rate of tenant database describes how the capacity of each tenant database increases as they continue to use it. Do you believe that this could be one of the factors affecting the adoption MTD?

1. Yes
2. No

32. How can you rank the degree of effect the growth rate of tenant database has on the adoption of MTD? Ranking from 1-5.

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

33. The combination of a Time factor and a Cost factor could be called ECONOMIC factor. What do you think is the degree of effect of this ECONOMIC factor has on the adoption of MTD? Ranking from 1-5

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

34. Does this Economic factor drive organisations towards adoption or rejection of MTD?

1. ADOPTION
2. REJECTION

35. The combination of the following factors data isolation, scalability, flexibility and customization will form what is called SECURITY factor. What do you think is the degree of effect this SECURITY factor has on the adoption of MTD? Ranking from 1-5

1. No influence -1

2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

36. Does this security factor drive organisation towards adoption or rejection of MTD?

1. ADOPTION
2. REJECTION

37. The following factors size of tenant database, number of tenants, number of users per tenant, Growth rate of number of tenants and growth rate of tenant database can be term as GROWTH factor. What is the degree of effect GROWTH has on the adoption of MTD?

1. No influence -1
2. Very Mild - 2
3. Mild - 3
4. Influential -4
5. Very Influential - 5

38. Does this growth factor drive organisations towards adoption or rejection of MTD?

1. ADOPTION
2. REJECTION

39. Do regulations like the Copyright and Rights in Databases Regulations 1997 as a factor drives organisation towards adoption or rejection of MTD?

1. ADOPTION
2. REJECTION

40. In your personal opinion, in the last 2 years, how has the acceptability or acceptances of multi-tenant database changed? Ranking from 1-5

1. Highly Reduced - 1
2. Reduced - 2
3. No difference - 3
4. Improved - 4
5. Highly Improved - 5

Appendix D - Cover letter for Validation

Faculty of Science and Engineering
University of Wolverhampton
Wulfruna Street
Wolverhampton
WV1 1LY
United Kindom
T. +44(0)1902 321000
F. +44(0)1902 321478

W. www.wlv.ac.uk/FSE

11TH April, 2016

Dear Sir/Madam,

VALIDATION OF A FRAMEWORK FOR ESTABLISHING A STANDARD SCIENTIFIC GUIDELINE FOR THE EVALUATION AND ADOPTION OF MULTI-TENANT DATABASE

Thank you for participating in the data gathering for my PhD research. I have now identified factors that can influence the adoption of Multi-Tenant Database (MTD) Management platform and the degree of impact of each of the factors. The research has been able to come up with a set of guidelines that can assist intending users of MTD in the process of evaluation and adoption of MTD. These guidelines have been developed into a framework and the framework into an expert system.

It is thought that these framework and expert system would be a useful resources to the government, MTD service providers and MTD users, particularly to enhance the decision making process during the adoption of the concept. In view of this, I would be very grateful if you could please respond to the feedback form, to help establish the relevance of the research findings and recommendations. As before, confidentiality and anonymity are guaranteed as all the information gathered will conform to the University's Ethical procedure. Please kindly complete feedback form through the link attached. Alternatively, if you wish to give the feedback over the telephone please send me an email. Contact details are provided below. Kindly look into the attached framework and the link to the expert system before filling the feedback form.

I would like to thank you in advance for your valued and kind consideration.

If you would like to receive further information about the research, please feel free to contact me.

Yours faithfully,

Olumuyiwa Matthew

Doctoral Research Student
Faculty of Science & Engineering
University of Wolverhampton
Wulfruna Street, Wolverhampton
WV1 1LY
Tel: [number redacted]
Mobile: [number redacted]
E-mail: [e-mail address redacted]

Appendix E – Questionnaire for Validation

Questionnaire for Validation of the Research Findings

Please provide response on how valid the research findings are with respect to your experience.

1. Cost in this research is the total cost of ownership which includes infrastructural cost, management cost and application cost, which is found to be a factor that encourages the adoption of Multi-tenant database (MTD). To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
2. Time in this research is the total number of days/hours/minutes it will take to build, configure and deploy a dedicated system to the site or premises of owner. This is found to be a factor that encourages the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
3. The combination of cost and time forms the economic factor, which is found to be a factor that encourages the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
4. Based on the findings, economic factor has the highest degree of impact towards adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
5. The size of tenant's database, the number of tenants, the number of users per tenant, the growth rate of number of tenants and the growth rate of tenant databases are all factors that are combined to form the growth factor. This is found to also encourage the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
6. This growth factor is found to have third highest impact value on MTD adoption. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
7. Data isolation, scalability, flexibility, and customization are all factors that are combined to form the security factor. This security factor is found to discourage the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
8. The security factor is found to have second impact value with a negative impact on the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
9. Regulations are laws and policies that govern and protect data. This regulation as a factor is found to encourage the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.

10. The regulation factor is found to have the least impact value among the four factors on the adoption of MTD. To what extent do you agree with this finding? Please tick [✓] one option.
    ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.

SUGGESTED RECOMMENDATIONS

Based on the research findings the following recommendations have been given. Please provide responses on the relevance of the recommendations with regards to your experience.

1. Experts in the field of Database management system should be given the responsibility of heading the project team when considering its adoption. To what extent do you agree with this recommendation? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
2. The economic factor must be considered, which includes the cost and time factor as explained in earlier sections. To what extent do you agree with this recommendation? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
3. The level of security that you want your MTD to achieve and tolerate should examine which include data isolation, scalability, flexibility and customization. To what extent do you agree with this recommendation? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
4. The growth rate of an MTD should also be considered in terms of the size of tenant database, number of tenants, number of users per tenant, growth rate of tenants and growth rate of tenant database. To what extent do you agree with this recommendation? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
5. The harmonious balance between the regulations governing prospective tenants should be look into by both the tenant and the service provider. To what extent do you agree with this recommendation? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
6. The choice of the MTD model must be determined by evaluating the features of the three model approaches of the concept. To what extent do you agree with this recommendation? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree.
7. Would you say that the framework is capable of assisting individuals and organisations in taking an informed decision about the adoption of MTD? Please tick [✓] one option.
   ( ) Not sure of its capability ( ) No, not capable ( ) Neutral ( ) Yes, capable ( ) Yes, highly capable
8. Would you say that the framework has incorporated all aspects needed for this decision making process? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree
9. Would you say that the expert system is simple and user friendly for intending tenants and does not require any amendment? Please tick [✓] one option.
   ( ) Strongly Disagree ( ) Disagree ( ) Neutral ( ) Agree ( ) Strongly Agree

Please provide any additional comments here (Please add extra pages if required).

Appendix F - Analysis of Validation Results using SPSS

```
FREQUENCIES VARIABLES=COST TIME ECONOMICS ECONIMPACT GROWTH GROWTHIMPACT
SECURITY SECURITYIMPACT REGULATION REGULATIONIMPACT EXPERTS
ECONCONSIDERATION LEVELOFSECURITY GROWTHRATE REGULATIONBALANCE
MTDmodelCHOICE FRAMEWORKCAPABILITY FRAMEWORKCOMPLETENESS
Esfriendliness
  /BARCHART FREQ
  /ORDER=ANALYSIS.
```

Frequencies

Frequency Table

COST

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	7	33.3	33.3	33.3
	Strongly Agree	14	66.7	66.7	100.0
	Total	21	100.0	100.0	

TIME

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	3	14.3	14.3	14.3
	Agree	9	42.9	42.9	57.1
	Strongly Agree	9	42.9	42.9	100.0
	Total	21	100.0	100.0	

ECONOMICS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	7	33.3	33.3	33.3
	Strongly Agree	14	66.7	66.7	100.0
	Total	21	100.0	100.0	

ECONIMPACT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	12	57.1	57.1	57.1
	Strongly Agree	9	42.9	42.9	100.0
	Total	21	100.0	100.0	

GROWTH

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	3	14.3	14.3	14.3
	Agree	9	42.9	42.9	57.1
	Strongly Agree	9	42.9	42.9	100.0
	Total	21	100.0	100.0	

GROWTHIMPACT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Disagree	1	4.8	4.8	4.8
	Neutral	6	28.6	28.6	33.3
	Agree	5	23.8	23.8	57.1
	Strongly Agree	9	42.9	42.9	100.0
	Total	21	100.0	100.0	

SECURITY

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	9	42.9	42.9	42.9
	Strongly Agree	12	57.1	57.1	100.0
	Total	21	100.0	100.0	

SECURITYIMPACT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Disagree	1	4.8	4.8	4.8
	Agree	9	42.9	42.9	47.6
	Strongly Agree	11	52.4	52.4	100.0
	Total	21	100.0	100.0	

REGULATION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Disagree	1	4.8	4.8	4.8
	Neutral	3	14.3	14.3	19.0

	Agree	7	33.3	33.3	52.4
	Strongly Agree	10	47.6	47.6	100.0
	Total	21	100.0	100.0	

REGULATIONIMPACT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	2	9.5	9.5	9.5
	Agree	11	52.4	52.4	61.9
	Strongly Agree	8	38.1	38.1	100.0
	Total	21	100.0	100.0	

EXPERTS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	1	4.8	4.8	4.8
	Agree	12	57.1	57.1	61.9
	Strongly Agree	8	38.1	38.1	100.0
	Total	21	100.0	100.0	

ECONCONSIDERATION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	1	4.8	4.8	4.8
	Agree	5	23.8	23.8	28.6
	Strongly Agree	15	71.4	71.4	100.0
	Total	21	100.0	100.0	

LEVELOFSECURITY

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	10	47.6	47.6	47.6
	Strongly Agree	11	52.4	52.4	100.0
	Total	21	100.0	100.0	

GROWTHRATE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	1	4.8	4.8	4.8
	Disagree	1	4.8	4.8	9.5
	Neutral	6	28.6	28.6	38.1
	Agree	8	38.1	38.1	76.2
	Strongly Agree	5	23.8	23.8	100.0
	Total	21	100.0	100.0	

REGULATIONBALANCE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Disagree	1	4.8	4.8	4.8
	Neutral	2	9.5	9.5	14.3
	Agree	8	38.1	38.1	52.4
	Strongly Agree	10	47.6	47.6	100.0
	Total	21	100.0	100.0	

MTDmodelCHOICE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	9	42.9	42.9	42.9
	Strongly Agree	12	57.1	57.1	100.0
	Total	21	100.0	100.0	

FRAMEWORKCAPABILITY

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	1	4.8	4.8	4.8
	Yes, Capable	13	61.9	61.9	66.7
	Yes, highly Capable	7	33.3	33.3	100.0
	Total	21	100.0	100.0	

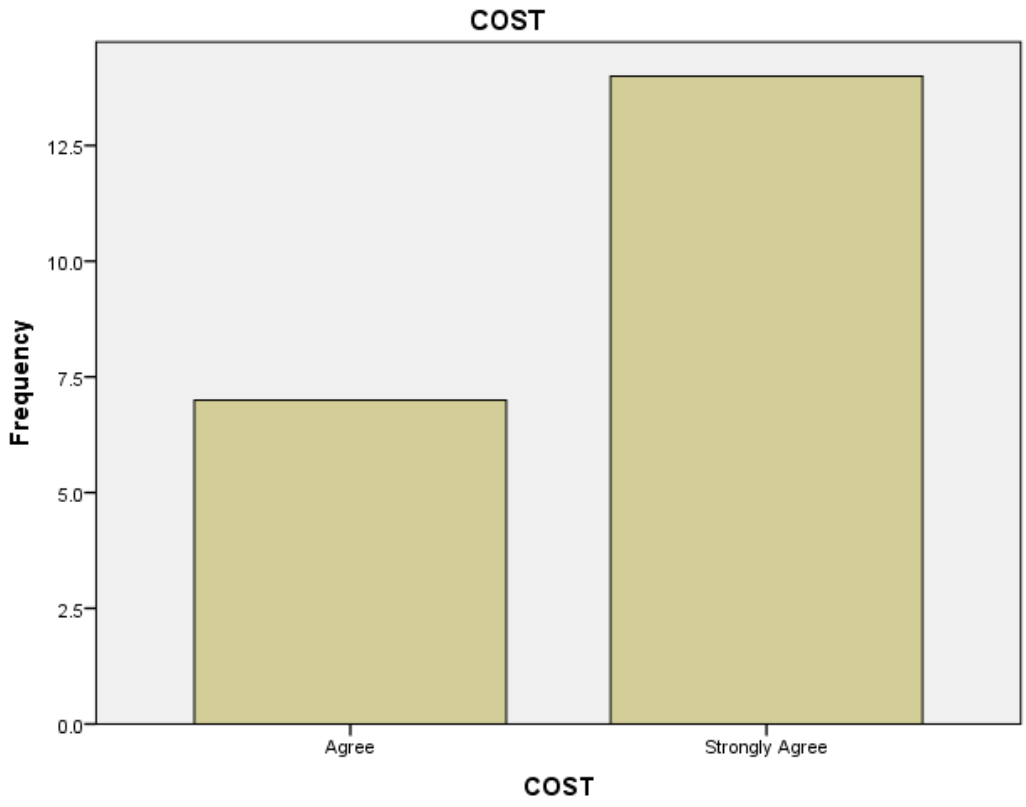
FRAMEWORKCOMPLETENESS

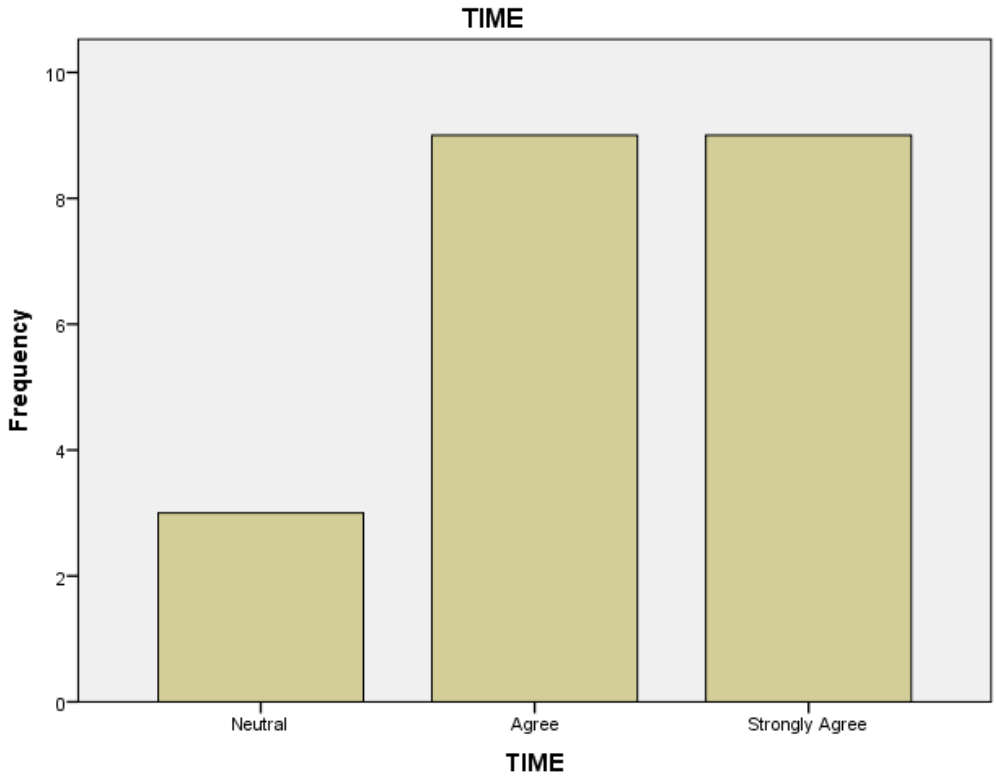
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Disagree	2	9.5	9.5	9.5
	Neutral	2	9.5	9.5	19.0
	Agree	11	52.4	52.4	71.4
	Strongly Agree	6	28.6	28.6	100.0
	Total	21	100.0	100.0	

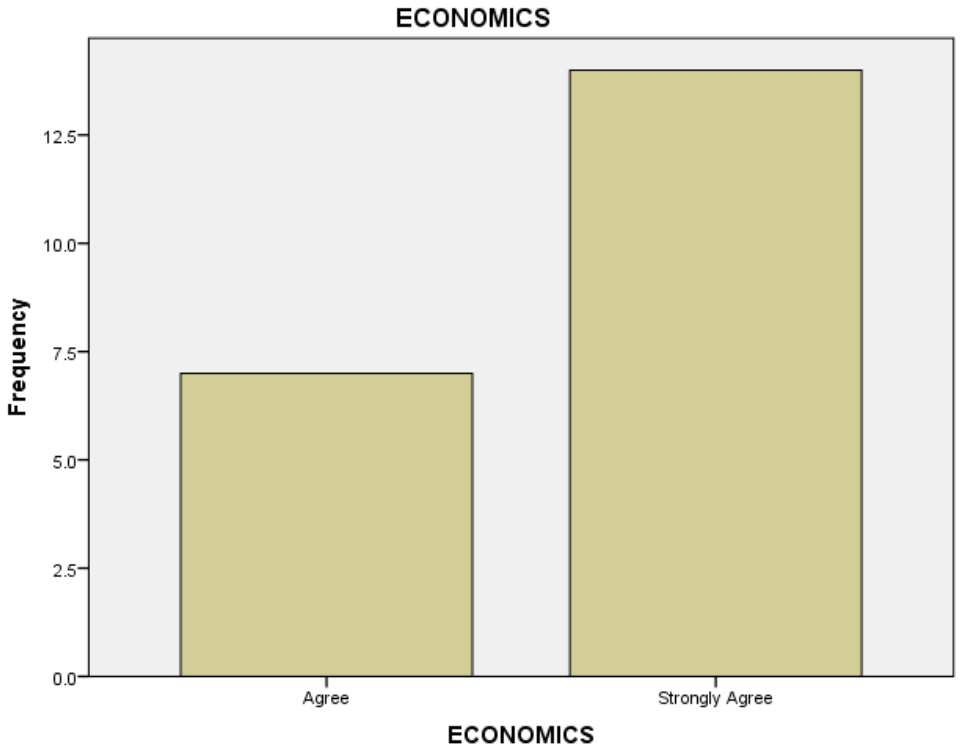
Esfriendliness

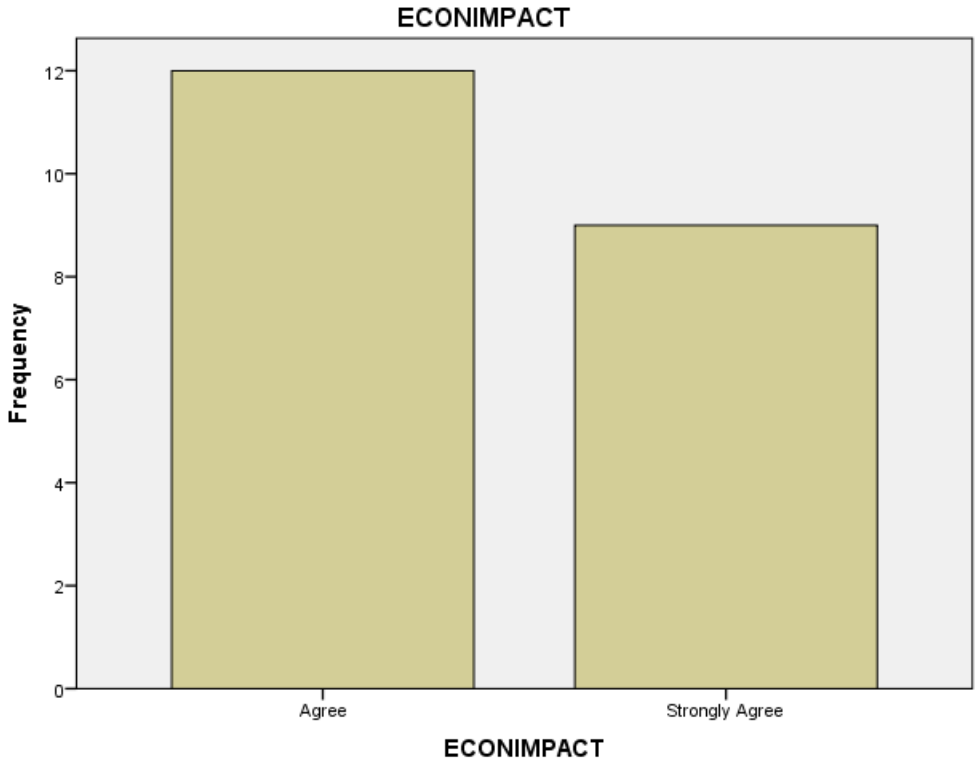
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Neutral	1	4.8	4.8	4.8
	Agree	9	42.9	42.9	47.6
	Strongly Agree	11	52.4	52.4	100.0
	Total	21	100.0	100.0	

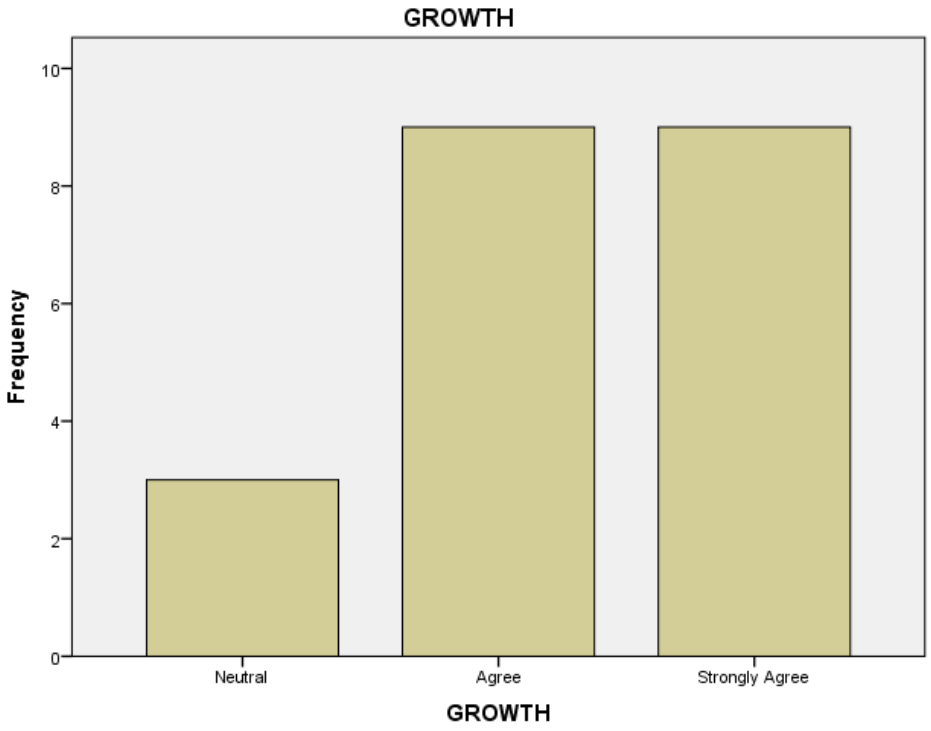
Bar Chart

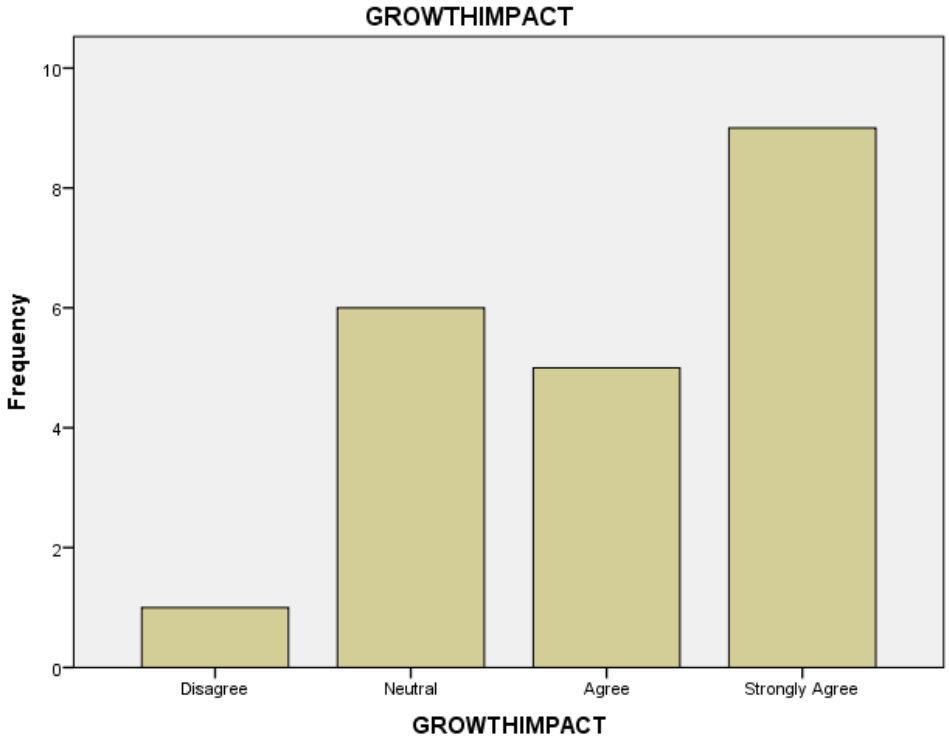


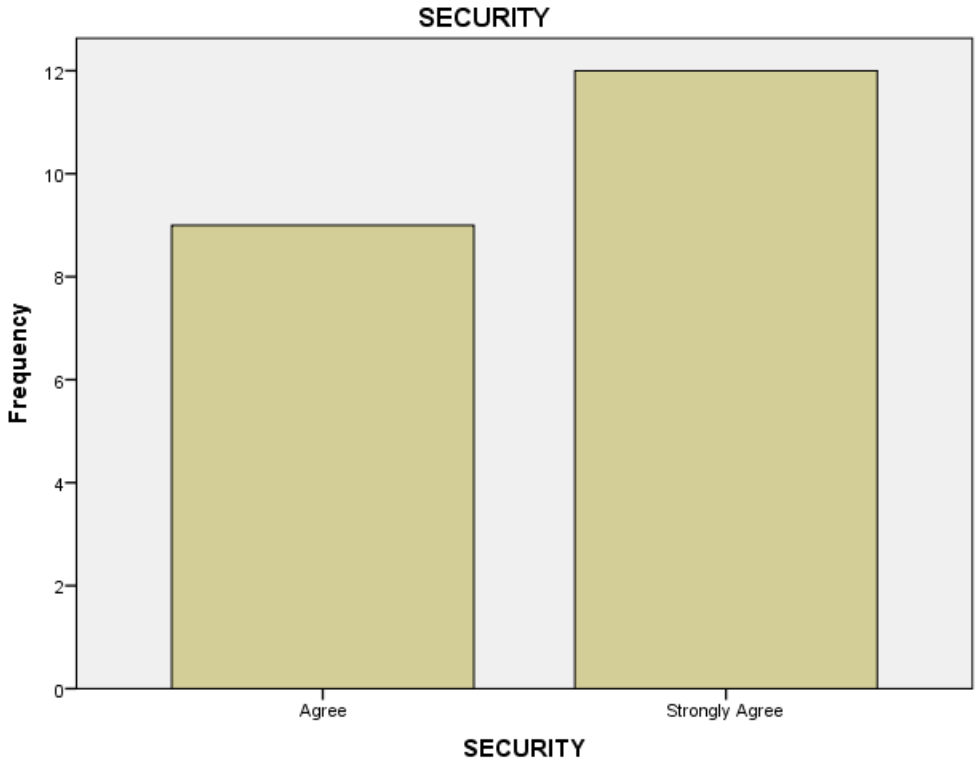

TIME
Frequency
10
8
6
4
2
0
Neutral
Agree
Strongly Agree
TIME

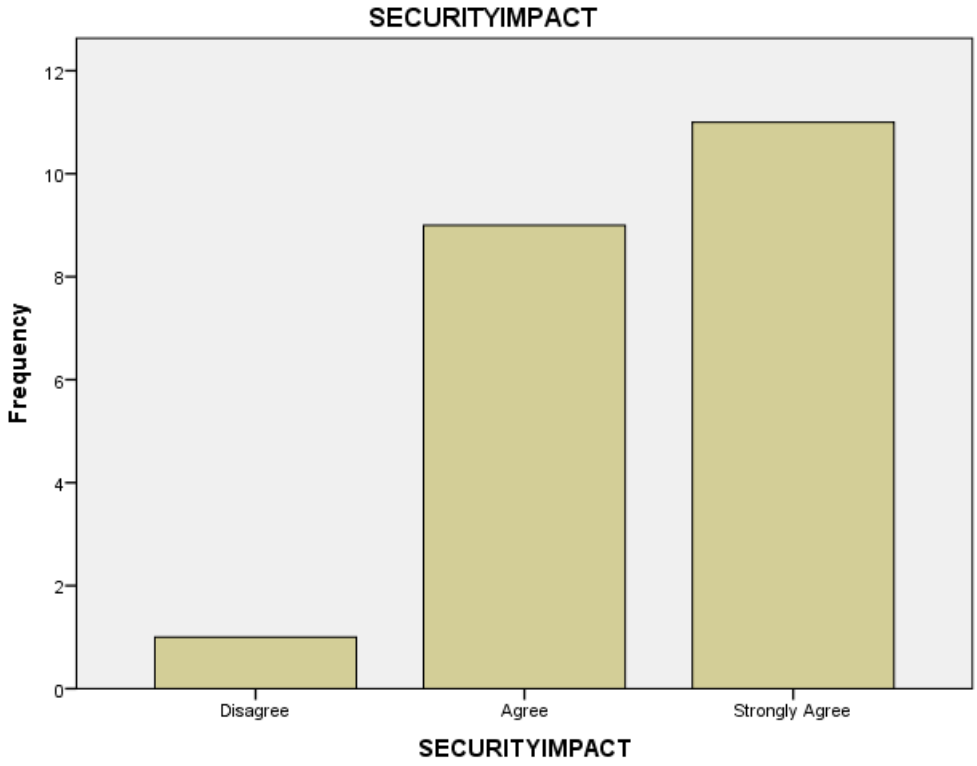

ECONOMICS
Frequency
12.5
10.0
7.5
5.0
2.5
0.0
Agree
Strongly Agree
ECONOMICS

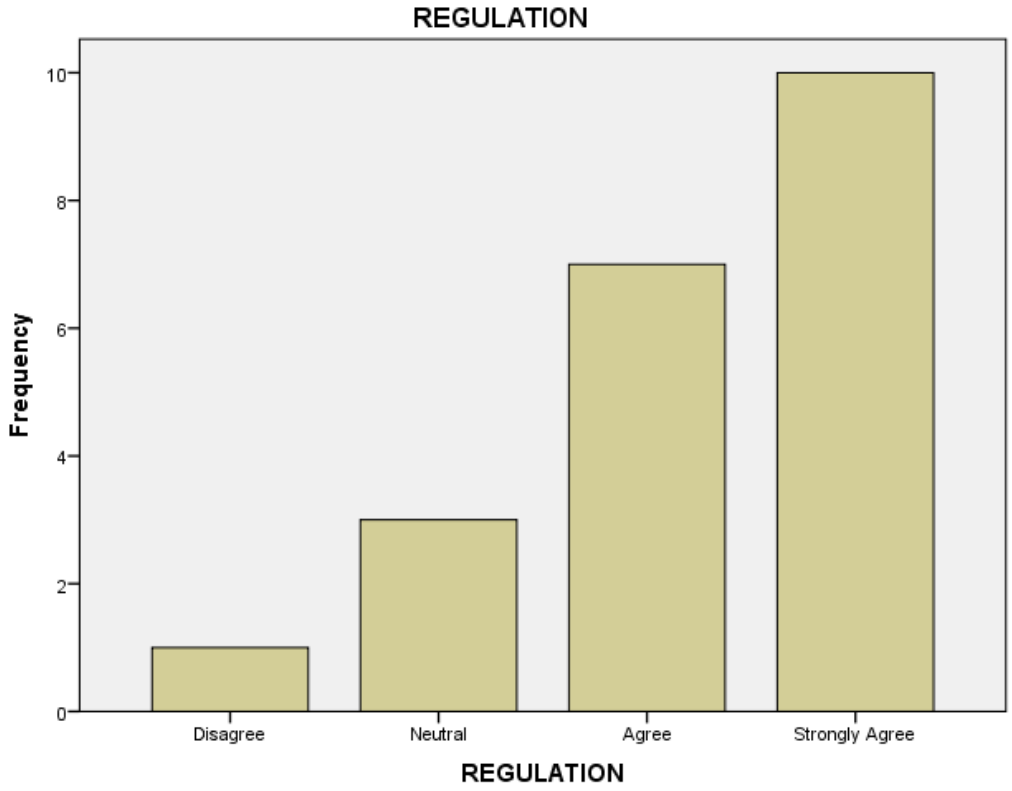

ECONIMPACT
Frequency
12
10
8
6
4
2
0
Agree
Strongly Agree
ECONIMPACT

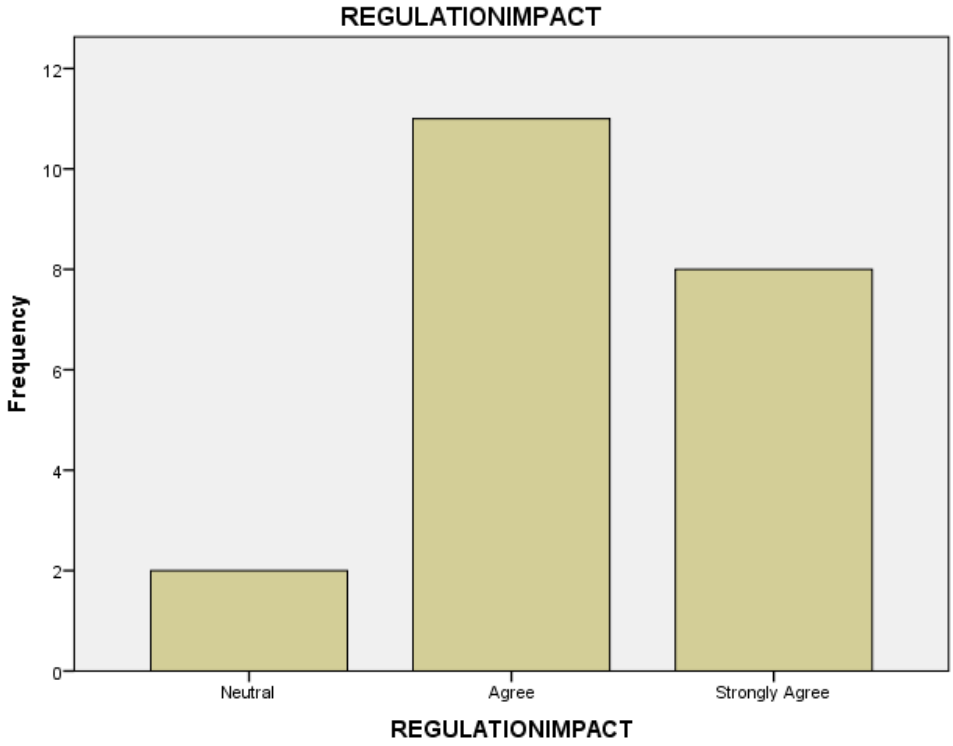

GROWTH
Frequency
10
8
6
4
2
0
Neutral
Agree
Strongly Agree
GROWTH

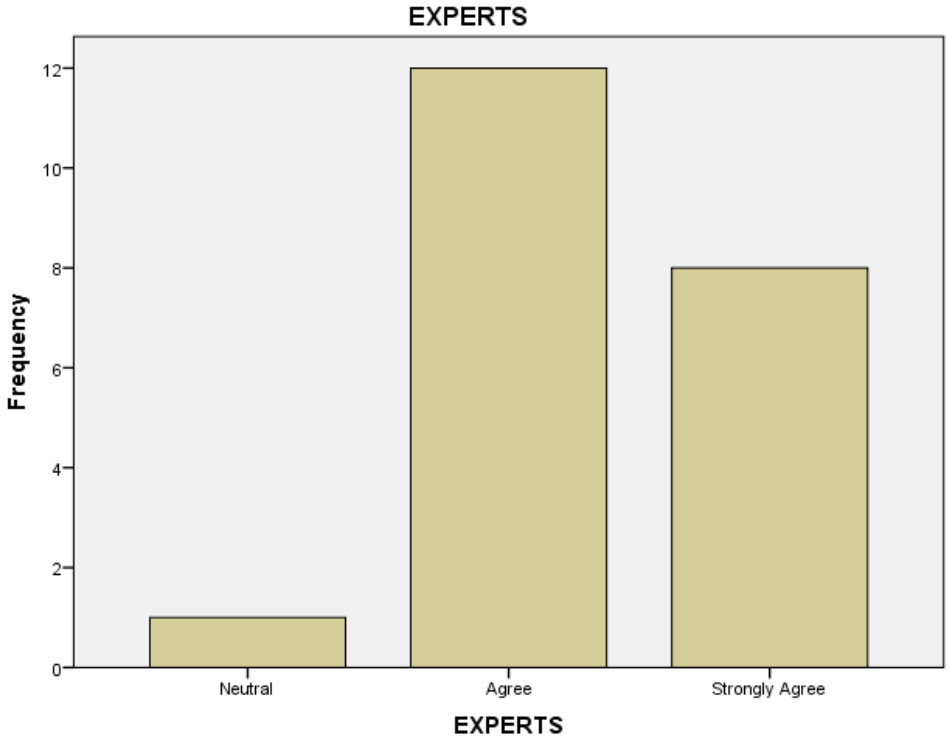

GROWTHIMPACT
Frequency
10
8
6
4
2
0
Disagree
Neutral
Agree
Strongly Agree
GROWTHIMPACT

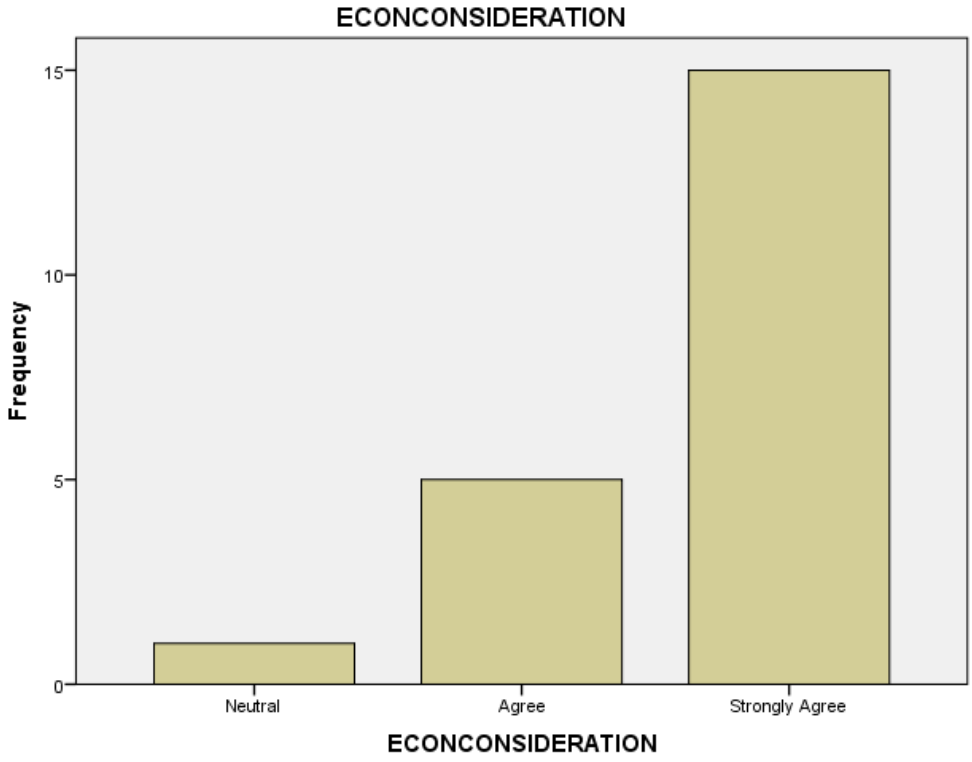

SECURITY
Frequency
12
10
8
6
4
2
0
Agree
Strongly Agree
SECURITY

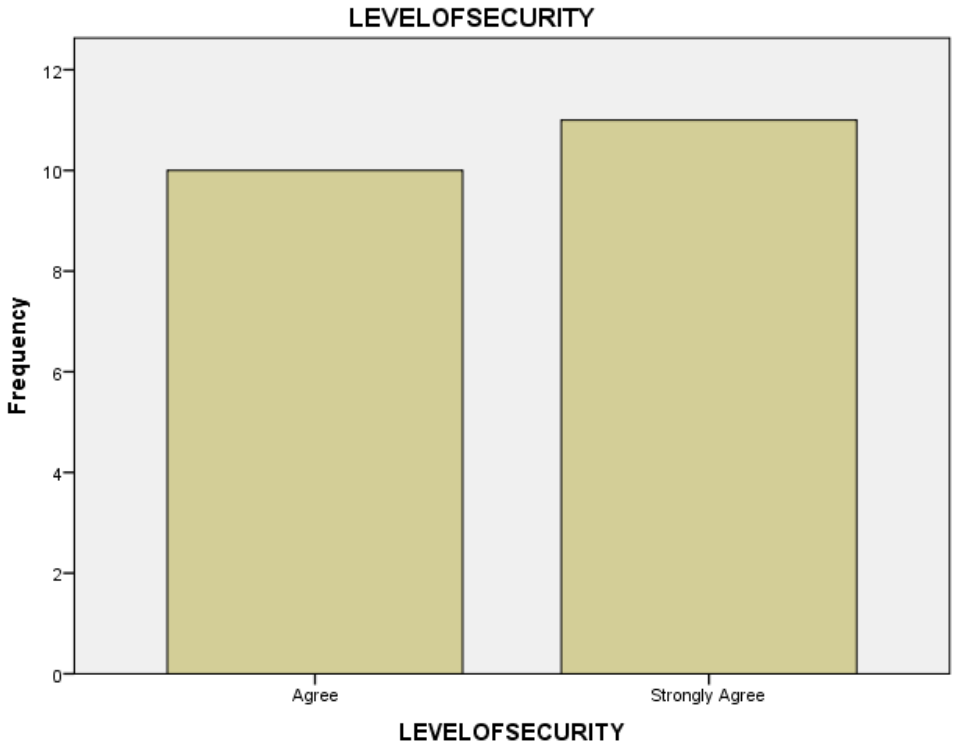

SECURITYIMPACT
Frequency
12
10
8
6
4
2
0
Disagree
Agree
Strongly Agree
SECURITYIMPACT

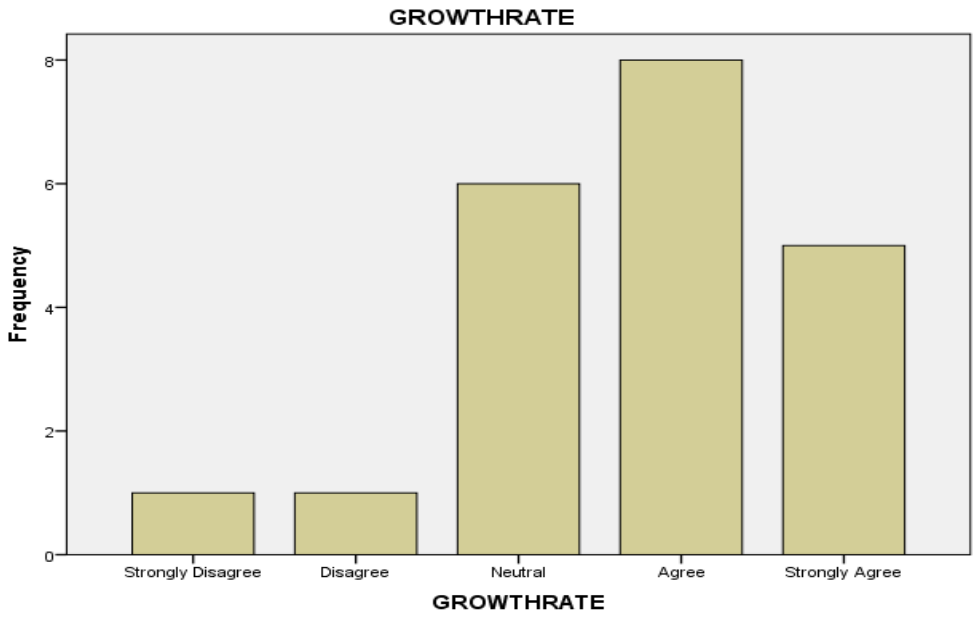

REGULATION
Frequency
10
8
6
4
2
0
Disagree
Neutral
Agree
Strongly Agree
REGULATION

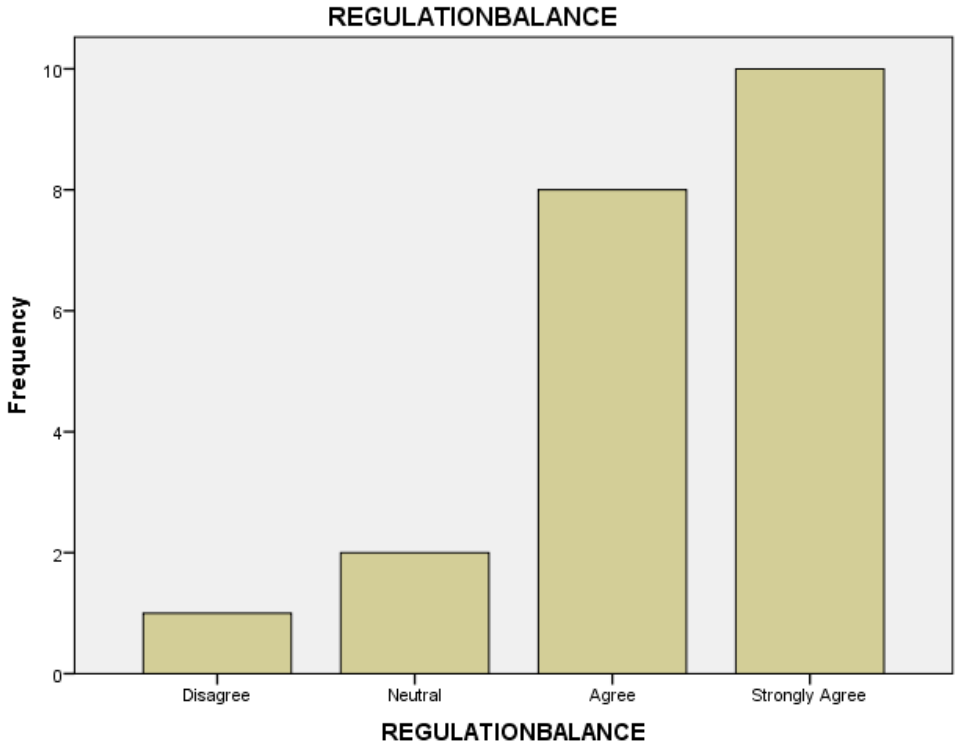

REGULATIONIMPACT
Frequency
12
10
8
6
4
2
0
Neutral
Agree
Strongly Agree
REGULATIONIMPACT

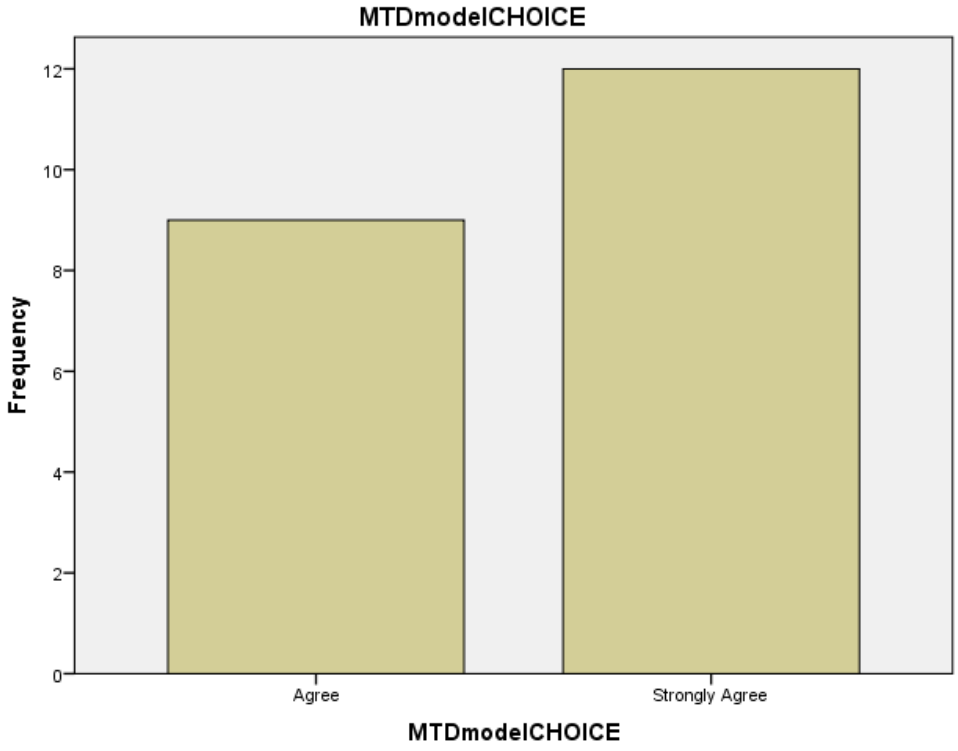

EXPERTS
Frequency
12
10
8
6
4
2
0
Neutral
Agree
Strongly Agree
EXPERTS

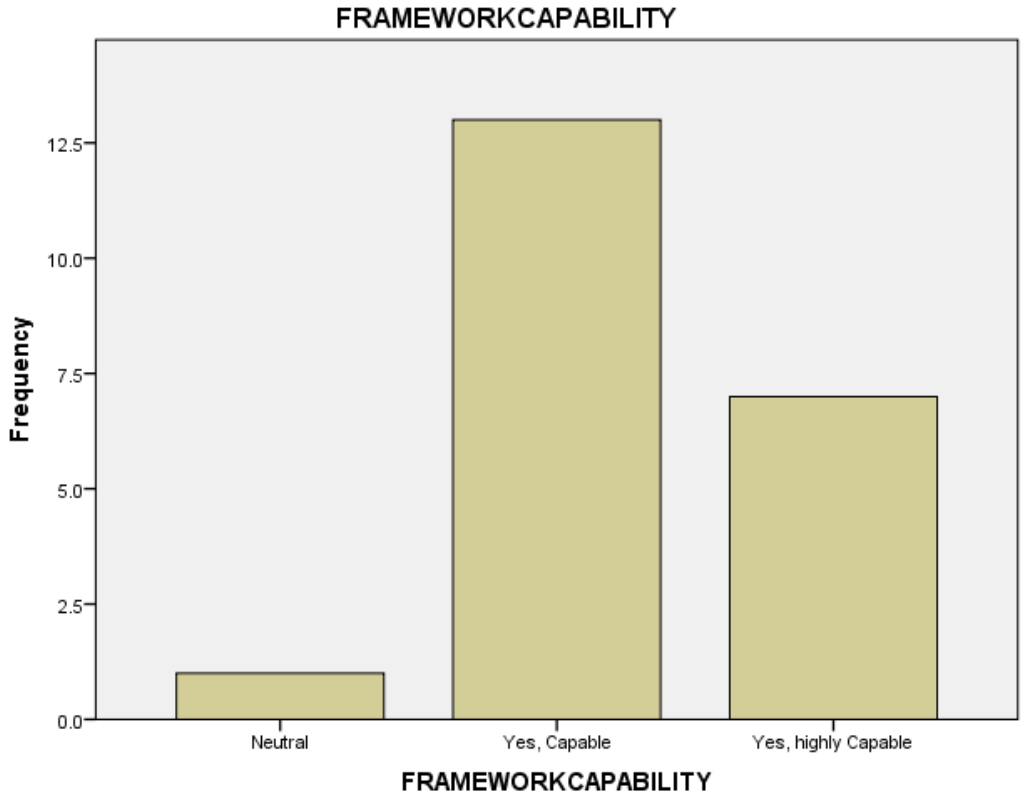

ECONCONSIDERATION
15
10
5
0
Frequency
Neutral
Agree
Strongly Agree
ECONCONSIDERATION

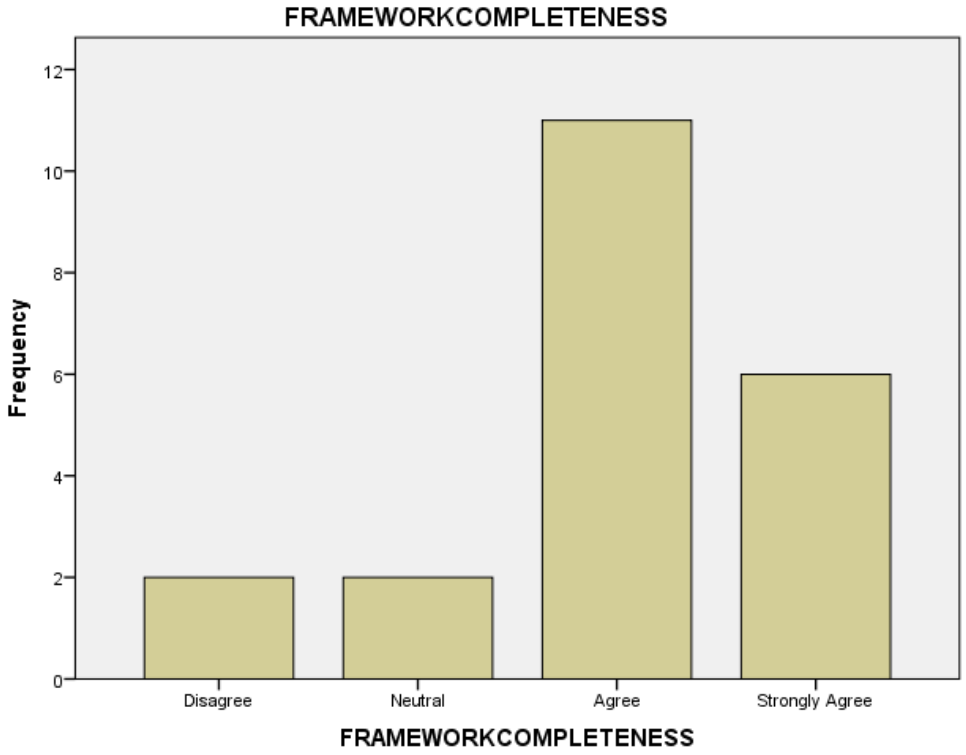

LEVELOFSECURITY
12
10
8
6
4
2
0
Frequency
Agree
Strongly Agree
LEVELOFSECURITY

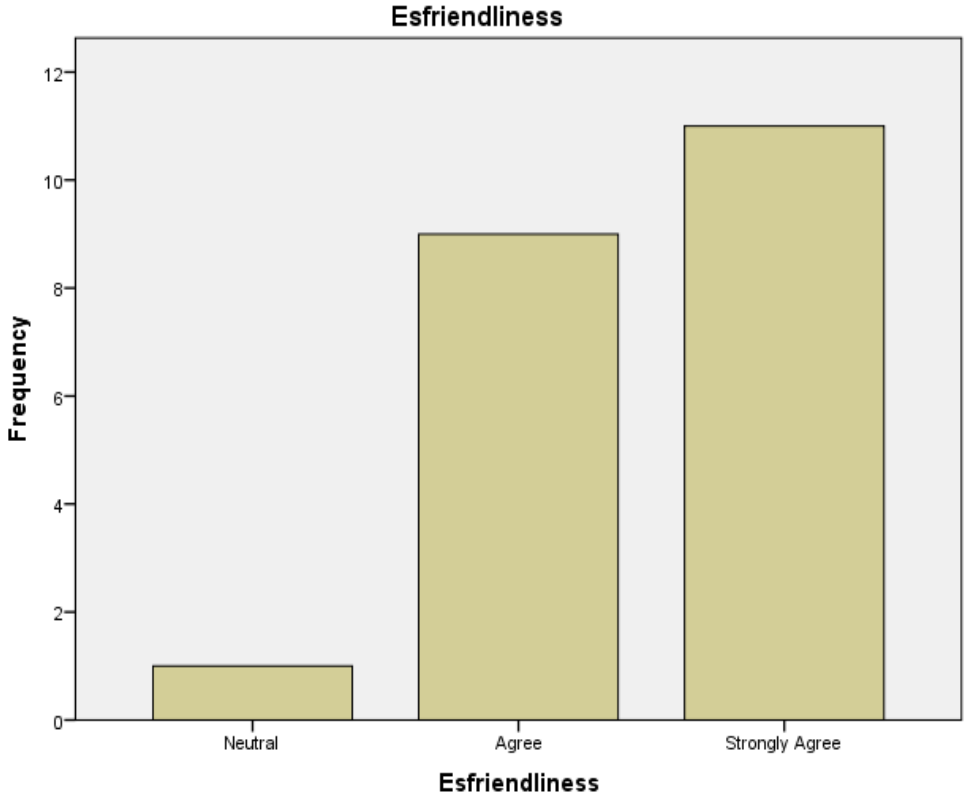

GROWTHRATE
Frequency
0
2
4
6
8
Strongly Disagree
Disagree
Neutral
Agree
Strongly Agree
GROWTHRATE


REGULATIONBALANCE
Frequency
0
2
4
6
8
10
Disagree
Neutral
Agree
Strongly Agree
REGULATIONBALANCE


MTDmodelCHOICE
Frequency
12
10
8
6
4
2
0
Agree
Strongly Agree
MTDmodelCHOICE


FRAMEWORKCAPABILITY
Frequency
12.5
10.0
7.5
5.0
2.5
0.0
Neutral
Yes, Capable
Yes, highly Capable
FRAMEWORKCAPABILITY


FRAMEWORKCOMPLETENESS
Frequency
12
10
8
6
4
2
0
Disagree
Neutral
Agree
Strongly Agree
FRAMEWORKCOMPLETENESS


Esfriendliness
Frequency
12
10
8
6
4
2
0
Neutral
Agree
Strongly Agree
Esfriendliness

Appendix G - Personal Reflection

The 45 months of the PhD research has been an intense learning period which has changed my attitude and personal life. There was always so much to do in a short time and it seemed to be a task that would never end. The research journey has taught me how to conduct independent study and has also stretched my intellectual capabilities. There are many things I have learnt along this thesis writing journey which include effective thesis management and scholarly writing. With the support from my supervisors, I now have a better understanding of the research process, building arguments, and writing academic reports. Overall, I have learnt valuable life lessons from the PhD process. Although I was faced with several physical, psychological, and emotional challenges, there were some good experiences especially with respect to further developing my intellectual capability.